2009 William Allan Award Address: Life in The Sandbox: Unfinished Business

Huntington F. Willard[1,*]

Thank you, Terry, for that introduction; and my deep thanks to the Awards Committee and to all of you for the award you have given me this afternoon. I am so grateful for and humbled by this honor. I am grateful because of your willingness to recognize the achievements of a large group of hardworking individuals who have worked and played in what we've called The Sandbox for the past 27 years. And humbled because I recognize that the list of people I now join is such a profoundly distinguished one. This is a wonderful group to now be part of, and I'm deeply grateful to The Society and to the Awards Committee for that honor. I'm so pleased to join my chromosome colleagues Pat Jacobs, Dorothy Warburton, and, of course, Mary Lyon, and a whole host of other individuals whom I've spent so much of my career looking up to. I'm also thrilled to share this afternoon with Janet Rowley, this year's winner of the Gruber Prize in Genetics. Between the Allan Award, the Gruber Prize, and the Nobel Prize for telomeres a week ago, it's been a pretty good month for chromosomes!

At the risk of giving away the punch line at the beginning, my comments this afternoon about the last 30-plus years of my scientific life will be less a personal tale of my own career than they are a tale of students and mentors, their achievements, their motivations, and the insights they've brought to their work. It's a story of young scientists who at the very early stages of their careers are empowered to take ownership of science, to take a question, to put their mark on it, and to decide what they want to do with it. Not what I wanted to do with it or what the field wanted to do with it, but what each of these students wanted to do with it—students validated in their search for trying to figure out answers to a series of questions that perhaps only they know, as they've articulated it (or perhaps haven't even yet fully articulated) to themselves.

Whatever successes I've enjoyed are most directly a reflection of the opportunities I was given very early on in my own training. If I've had success since starting my own group, it's been success in creating an environment in which students could come into The Sandbox—whether graduate students, postdoctoral fellows, or, now increasingly, undergraduates—and could find an opportunity to be validated as scientists, to realize that a lifetime of discovery is one of the greatest privileges that one can have in science. Such an awakening is not something that comes naturally to many students, because there's nothing in our general society and certainly little in our educational system in its earliest stages that prepares students for the recognition of what a life of science and a life of discovery is all about and for the sense of empowerment that comes with that.

The Beginning

So this is a story of students, my students. But I start with a story of another student some 36 years ago. This story begins on October 31, 1973, almost exactly 36 years ago. Now, you might say, "How did he ever remember that that was the date?" Well, I know that was the date because I actually keep my notebooks! I was taking Biology 113 then in college, a course on human genetics that used the "red book," the textbook by Curt Stern that remains, in my view, the best textbook on human genetics that has ever been written and certainly the book that opened this field up to me.[1] On October 31, 1973, we had a lecture on X inactivation. And as I took my notes that day (Figure 1)—and believe me, my handwriting was a whole lot better 36 years ago than it is now!—I began to realize that this was an unbelievably interesting and what I would come to call a "chewy" problem—a problem that had absolutely no precedent to suggest an answer, no set of guidelines or rules that we could understand at that time or that we could use to even begin to think about the problem.

This was a wonderful time in human genetics, as those of you who are of that vintage will recognize. Mary Lyon herself had only written about X inactivation a dozen years before this,[2] and this was a period of time, especially in human cytogenetics, during which the community of human and medical geneticists was trying to figure out just how relevant the idea of X chromosome inactivation might be to our field. As seen through the eyes of an impressionable young student, several spectacularly insightful and inviting papers were published at that time, gray and dusty copies of which I still lug around in my now heavily dog-eared file on X inactivation. Brown and Chandra's idea for how the X inactivation center works[3] still remains one of the most lucid models of

[1]Duke Institute for Genome Sciences & Policy, Duke University, Durham, NC 27708, USA
*Correspondence: hunt.willard@duke.edu
DOI 10.1016/j.ajhg.2010.01.037.

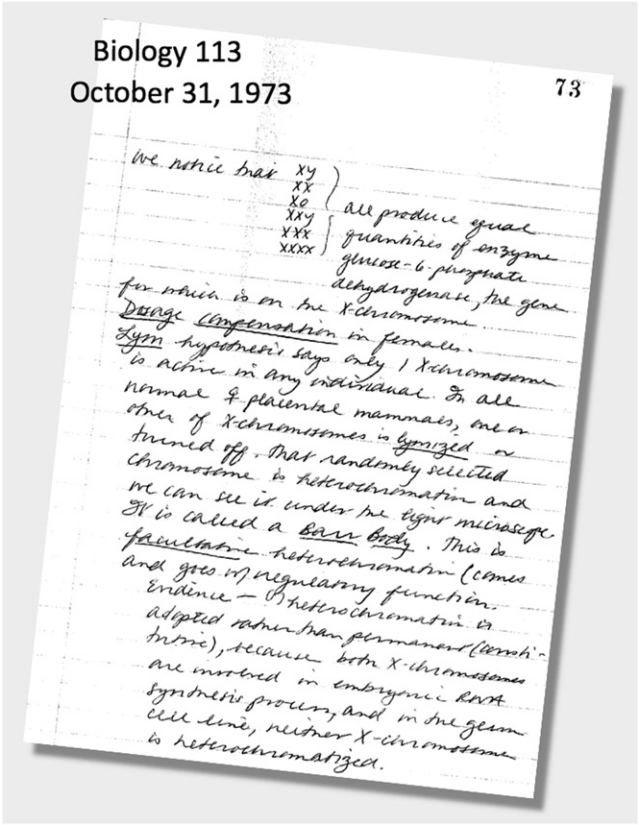

Figure 1. Notebook Entry from October 31, 1973
An impressionable student's first introduction to X chromosome inactivation.

random X inactivation, and several reviews by Eva Therman and Klaus Patau[4,5] served as both an introduction and the bridge that would bring X inactivation and chromosomes together for me.

So how was a young student going to get started? At that stage, and this was now early spring of 1974, I did what I thought all students would do, although I realize in retrospect that that wasn't necessarily true. I just picked up the phone and called the head of the Clinical Genetics Division at Children's Hospital and asked to speak to Park Gerald. Can you imagine? When that happens to me now and students call or email me, I try to remember that day in 1974 and try to be as generous with my time as Park was with me. That very day, he came on the phone and wanted to know who I was. Without hesitation, I asked if there were any opportunities to work in a laboratory that summer, because human genetics was going to be my future. Well, he deigned to meet with me a few days later, and after chatting with me for about 10 or 15 minutes, he said "Wait here," and he went down the hall to grab Sam Latt. Sam had published a paper[6] just a few months before, his first paper laying out the principles of using bromodeoxyuridine (BrdU) as a measure for the timing of replication based on the interactions between BrdU and the dye 33258 Hoechst, a method that revolutionized the study of chromosomes and allowed fluorescence detection of DNA replication instead of the hideously painful technique of autoradiography, which was in use up to that time.

I joined Sam's lab that afternoon and spent the next 18 months working under his tutelage at an incredibly exciting time when, although I didn't realize it at the time, I was allowed to be at the breaking edge of a wave, working with him trying to figure out the patterns of X inactivation and DNA replication on the active and inactive X chromosomes in female cells, at a level of resolution and with a level of precision that just simply hadn't been possible previously. Under very different and far more difficult circumstances, I've written elsewhere[7] about those times in his lab, but the thoughts are just as meaningful and relevant today:

> One of the potentially most rewarding by-products of our business is the strong, mutually supportive relationship that often develops between mentor and student. Some are personal, others intellectual; in some cases, the distinction blurs. In the most favorable instances, the mentor provides leadership, guidance, and inspiration, in return for satisfaction from setting a young colleague on the "right" path. For his [or her] part, the student benefits from the opportunity, when done right, to learn excellence and passion at a time when one's the most impressionable. The first teacher-student relationship has a unique flavor. Like other first encounters, it cannot be repeated, and its lessons – either good or bad – are most apt to be lasting ones…Since I had my first exposure to research in Sam's lab at a time of great excitement in cytogenetics, it is difficult to know how things might have turned out if he had not been willing to take a chance on a very green undergraduate and to share his sense of passion for chromosomes and discovery. I remember hours spent in the pitch darkness of the darkroom, waiting for autoradiographic emulsion to dry on dipped slides, listening to Sam go on about the latest in chromosome structure, X inactivation, or clinical cytogenetics. For a wide-eyed student fumbling in the dark, those sessions were profoundly stimulating. The experience of working with Sam has…had a lasting influence on the directions of our research and our thinking.

Sam and I published a paper early in 1976, my first paper in *The American Journal of Human Genetics*, and it remains one of the papers of which I'm most proud.[8] Figure 2 illustrates these active X and inactive X replication patterns, from one of the slides that I presented in my ten-minute talk at the October 1975 meeting of this Society, the first of my ASHG meetings. But I show it here again because there are parts of this story that, even some 35 years later, we still don't understand; every time I look at that picture of DNA replication on the active and inactive X chromosomes in female cells, I am struck by the size of the

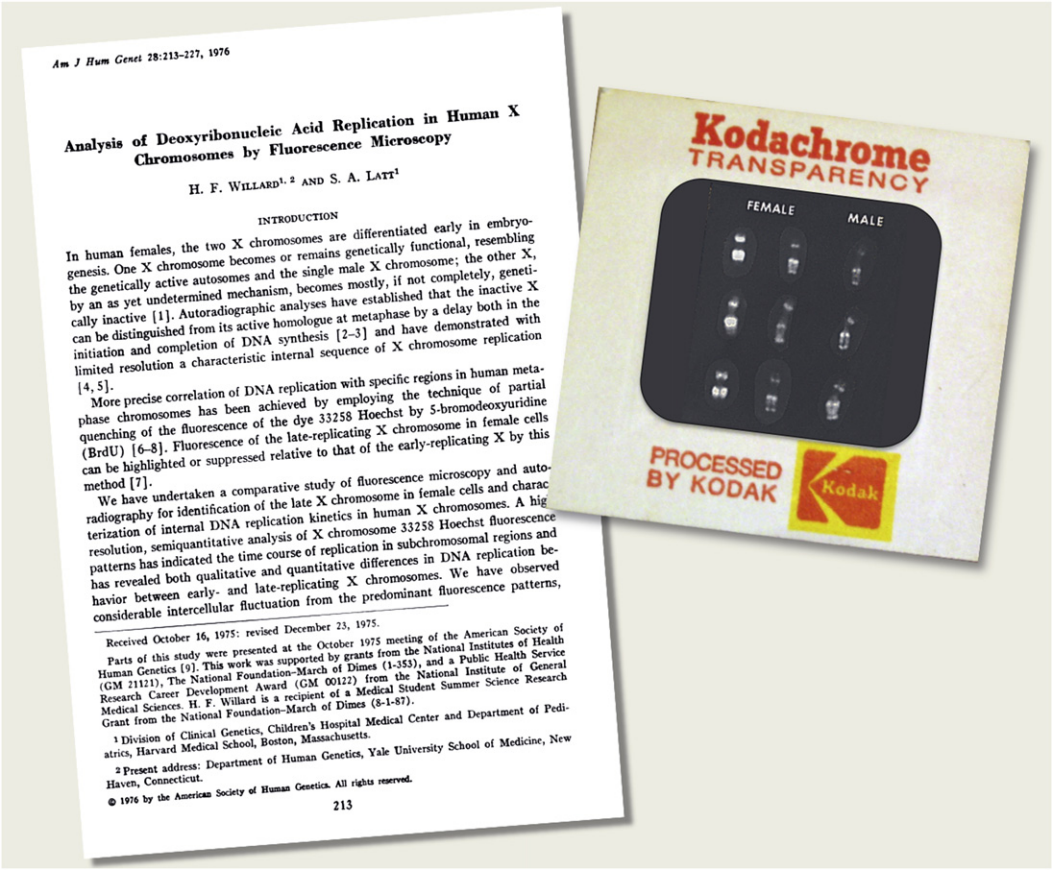

Figure 2. DNA Replication Patterns on Active and Inactive X Chromosomes
Left: first page of Willard and Latt,[8] published in *The American Journal of Human Genetics* in 1976. Right: slide of male and female X chromosome replication patterns, shown at the October 1975 meeting of ASHG in Baltimore, MD.

segments of the X chromosome—some 20 to 30 Mb of DNA—that coordinately shift their time of replication in S phase. We still don't have a clue in 2009 how that happens! But hopefully, with the ENCODE (Encyclopedia of DNA Elements) project and an increasing technological capability for studying questions like this, we will begin to understand what features of our genome and chromosomes control the timing of replication and, as a consequence of that, begin to understand what it is about the epigenetic control of X inactivation that so fundamentally shifts the replication behavior of the inactive X chromosome. Some still unfinished business.

As I completed my undergraduate work with Sam, and with my interest in human genetics, I searched around the country for graduate programs that would allow me to do what I wanted to do, which was, of course, to continue to study X inactivation. After a brief flirtation with Stan Gartler that almost took me to the west coast, I ended up going to Yale for a four-year side trip into inborn errors of metabolism, studying with Leon Rosenberg. While I recall being disappointed that each new disorder in cobalamin (vitamin B12) metabolism being defined then turned out to be autosomal and not X-linked, this was a wonderful time for me because, even though my Ph.D. projects weren't directly connected to X inactivation, it was exposure to biochemical genetics that has allowed me to appreciate much more deeply the rich nuances of the field of human genetics and to think much more broadly about genetics than I would have if I had stayed glued to chromosomes for my entire career.

At Yale, I benefitted enormously from the mentorship of Lee Rosenberg. As I commented in my ASHG Presidential Address eight years ago,[9] there's nothing in my career that I've achieved that I don't ultimately give him credit for. In reality, I actually had two mentors at Yale, one being Lee and the other Roy Breg, who directed the clinical cytogenetics lab at Yale at that time. It was Roy who allowed me to continue to work on X inactivation late in the evenings, and we were able to publish a few papers on that work at that time,[10,11] allowing me to keep my fingers connected to the X chromosome. Roy was important to me for another reason too, as he had the great wisdom to hire as a cytogeneticist in his lab Vicki Powers, my wife now of some 30 years. It was love at first sight when I met Vicki, because she had what every 25-year-old, red-blooded American male would want—she had a microscope, which I desperately needed to carry out my studies! So, that was a match made in heaven, and the rest is history.

The Sandbox and the X Chromosome

In January 1982, I took up a faculty position at the University of Toronto, and The Sandbox officially opened (although, I now realize, no one can actually remember when we started calling it that!), with the goal of tackling X inactivation. My office in Toronto from the very early days had a cartoon on the wall that I've always loved, because it represented then (and to a certain extent even now) what I felt we knew about X inactivation (Figure 3).

In starting out the lab at that time—and remember, this was some two decades before we'd have the human genome sequence—I developed a plan (hatched, as I recall, in discussions with Lee Rosenberg, walking on the shore near his house on Long Island Sound) to isolate DNA sequences from the X chromosome and, in my way of thinking, to use that to identify and clone genes that either were subject to or escaped from X inactivation. It was then that the very first examples of genes escaping from inactivation were becoming known, and these seemed to represent a promising avenue for understanding the chromosomal basis for X inactivation. What seemed then like a very straightforward plan turned into a much more convoluted but exciting path that has taken us into many different aspects of X inactivation, chromosome biology, and medical genetics over the ensuing 27 years (Figure 4).

The early strategy to develop approaches for mapping cloned DNA and studying the expression of genes along the chromosome grew into a series of studies that are still being carried on today in our group with ever-changing technologies to derive X inactivation profiles in females for different genes that escape or are subject to inactivation. This led us more deeply into the genomics of the X chromosome, into sex chromosome evolution, and into thinking about the different medical consequences of genetic imbalance for either short-arm or long-arm abnormalities of the X chromosome. It also allowed us more recently to gain insight into the extraordinary amount of variability that exists between females in the

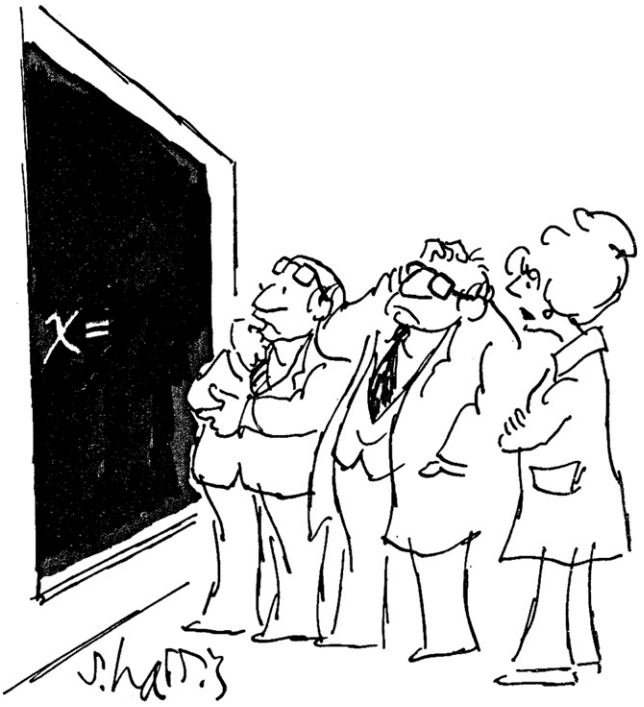

Figure 3. Inspiration for The Sandbox
Reprinted with permission from Sidney Harris. (Copyright sciencecartoonsplus.com.)

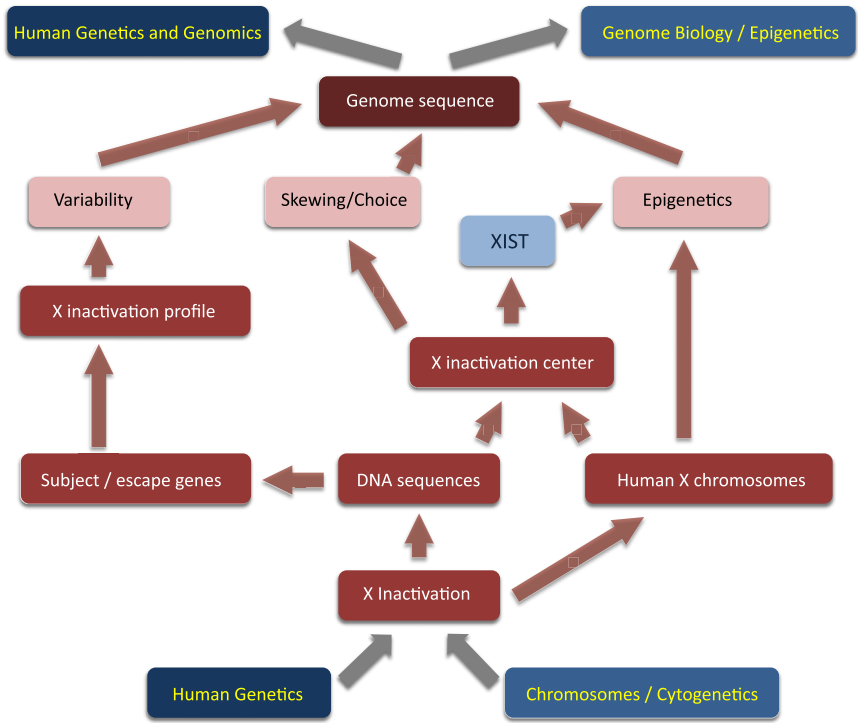

Figure 4. Intellectual Flow Diagram: X Chromosome Inactivation

population with respect to X-linked gene expression, the genetic and genomic basis of which we're still trying to understand.

Many of the studies of the X chromosome in The Sandbox began with Carolyn Brown, who joined my lab as a graduate student in those early years in Toronto. It was she who had the original insight to recognize a gene that escaped X inactivation in a location that suggested that there could be many more such genes up and down the length of the chromosome, not just in the pseudoautosomal region or even in the ancient pseudoautosomal region.[12] She later was instrumental in recruiting Laura Carrel to the lab, and together they came up with a much larger number of genes that escaped inactivation not only on the proximal short arm but also on the long arm of the X.[13] Carolyn and Laura brought great energy and persistence to what has turned into a multistudent and multidecade project, realizing over time that it would take 1200 genes, not just 12, to finally establish the genomic, chromosomal, and evolutionary basis for these X inactivation profiles. Yet more unfinished business!

As a graduate student at Stanford, Laura took this project on and established what was then the first-generation X inactivation profile.[14] Andy Miller, another Stanford graduate student, defined regions on the X chromosome in which there were multiple genes clustered in a domain of several hundred kilobases in which all of the genes escaped inactivation.[15] This clustering suggested that there was something about the genomic sequence or organization of the chromosome itself that was controlling or contributing to the epigenetic result of X inactivation or escape from X inactivation. Before she was done, Laura ended up analyzing some nearly 800 genes on the chromosome[16] in work that was published at the same time as the 155 Mb sequence of the X chromosome just a few years ago.[17]

The other part of the story for X inactivation, of course, is not just profiling X-linked gene expression, but determining the identity and nature of the X inactivation center, one of the big questions raised initially by Mary Lyon during the early 1960s.[18] As originally outlined by Eva Therman in some of the early papers[4,5] I read as an undergraduate, it was clear that human cytogenetic material would be critical for helping the field identify where the X inactivation center was and how it might function. After a number of years of finding and characterizing abnormal X chromosomes, this line of reasoning led to the identification of the *XIST* gene, another story that began with Carolyn Brown.

Working with Jim Rupert and Ron Lafreniere and then later with Brian Hendrich, another graduate student in the lab, Carolyn championed our early studies of the X inactivation center region[19] and pursued a positional cloning strategy, using candidate DNA or cDNA clones from various regions of the X chromosome. These candidate clones had been sent to us by helpful X chromosome colleagues around the world, most especially in this case including Andrea Ballabio (then at Baylor College of Medicine), who provided the initial candidate cDNA and whose group helped with the initial characterization of *XIST*, once we had found it. Carolyn and her team identified and fully explored *XIST*'s eponymous feature,[20,21] its inactive-X-specific transcription, whether that inactive X was found in normal females or in males carrying more than just a single X, as in Klinefelter syndrome.[20]

Carolyn had thus identified a gene expressed only from the inactive X and lying in the genetically and cytogenetically defined X inactivation center region, an obvious and strong candidate for a functional component of the X inactivation center itself. But how might it work? As she and I worked on the initial manuscripts to get them submitted to *Nature* in October of 1990 prior to her revealing the data for the first time at a workshop at the ASHG meeting that fall, we realized that a *cis*-acting functional RNA on the inactive X would be a more likely model than a protein product. Despite not having yet cloned or sequenced the full *XIST* RNA product (which, after all, was some 17 kb long), we proposed "that the *XIST* product is a *cis*-acting RNA molecule, perhaps involved structurally in formation of the heterochromatic Barr body."[20] This was the first long noncoding RNA to be implicated as having a particular function in human cells and, together with Shirley Tilghman's discovery of the H19 RNA,[22] provided early examples of the role of noncoding RNAs in epigenetic regulation.

Having the *XIST* gene in hand and understanding the beginnings of the X inactivation center enabled us to pursue two different kinds of studies: first, studying skewed patterns of inactivation in females, different ratios that could differ significantly from the random 50:50 "coin flip" expected from the Lyon Hypothesis, and second, exploring the epigenetic control of gene expression on the X, studies that have moved into the field of chromatin and epigenetics, focusing on noncoding RNAs as well as histone variants and modifications. It was Jim Amos-Landgraf and Robert Plenge who studied patterns of skewed inactivation in different cohorts of females, either normal females[23,24] or carriers of various X-linked clinical conditions.[25] It was Brian Chadwick and Cory Valley who began to explore chromatin aspects of X inactivation,[26–28] projects that continue in the lab today. Figure 5 illustrates the striking banding pattern of heterochromatin on the inactive X, which looks so reminiscent of the DNA replication timing patterns that Sam Latt and I had seen some 25 years before. Here, we're looking not at DNA replication timing, but rather chromatin variants and histone modifications on the inactive X, bringing nearly full circle the notion that the X chromosome—and by extension the whole genome—must be organized in a way that reflects a very careful interplay between the underlying sequence and the behavior, at a very large scale, of those sequences in the context of chromatin and the chromosome. The answer to that question is a life's pursuit and remains very much unfinished business.

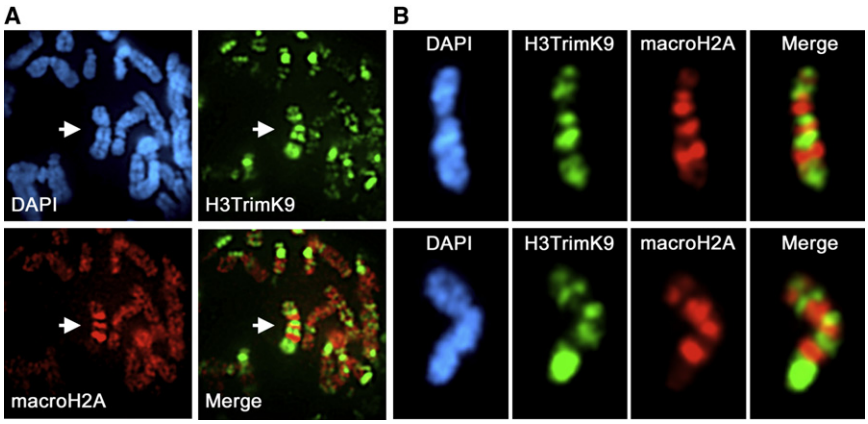

Figure 5. Epigenetic Patterns on the Inactive X Chromosome
(A) Indirect immunofluorescence staining of female metaphase chromosomes with antibodies to trimethylated form of histone H3 at lysine-9 (H3TrimK9) or to the histone variant macroH2A.
(B) Composite of inactive X chromosomes from two cells. Note banding pattern of alternating chromatin types,[27] reminiscent of patterns of X chromosome DNA replication.[8] (Images from Ph.D. thesis of C. Valley, Duke University, 2007.)

So, to me, the storyline shown in Figure 4 is a very straightforward one, reflecting a series of what seemed at the time to be logical and linear connections that have emerged in the decades since that day in October 1973. To you, you might be saying, "This was a plan?" But such is the nature of discovery science in human genetics (and certainly in human cytogenetics), in which so many roads, whatever the nature of the original motivation or insight, now seem to lead to genome sequence and to genome biology.

The Sandbox and Centromeres

But I haven't told you quite the whole story. If I go back to the very early days, when we were beginning to look at DNA sequences on the X in the late 1970s, it wasn't so much to clone genes that escaped from or were subject to inactivation. Rather, my initial hypothesis, when I was a postdoc working at Johns Hopkins with Kirby Smith and Barbara Schmeckpeper, was to identify an X-specific repeated DNA element that I thought, based on X inactivation models in the literature and based on the success that Lou Kunkel had had earlier in identifying tandemly repeated DNA from the Y chromosome,[29] would be blocks of X inactivation elements distributed along the chromosome,[30] similar to what has been hypothesized by Gartler and Riggs[31] as "way stations" or "booster elements" that could perpetuate the X inactivation signal up and down the length of the chromosome. And what a fine hypothesis that was! Other than, of course, the small fact that it seemed to be wrong…

My efforts as a postdoc to identify X-specific repeating DNAs involved in X inactivation turned out to be a failed experiment of sorts, designed initially as an experiment to study X inactivation that turned out to be anything but that. The reality was that what I isolated was X-specific centromeric DNA, which, once we realized it,[32] opened up a totally different side of the laboratory, a side that has allowed us to bridge back and forth for the past 25 years between two very different aspects of chromosome structure and function, one closely tied to gene expression, the other closely tied to chromosome biology.

Having identified centromeric DNA allowed us, in experiments championed initially by John Waye, to explore alpha satellite DNA[33] (a then recently discovered but poorly understood DNA family in the human genome), which took us into the study of a large number of chromosome-specific DNAs and their utilization, genomic mapping and the organization of those sequences, satellite DNA and repetitive DNA evolution, and finally centromere function, all of which came from my initial failed experiment in X inactivation (Figure 6). Somewhere there's a lesson in that particular story…

The potential significance of this change in direction became apparent in 1985 with another paper published in *The American Journal of Human Genetics*,[34] probably the first and certainly the last paper from The Sandbox in which all of the experiments were done with my own fingers. (It was shortly after this that the lab realized I was far better off sitting in my office and leaving the experiments up to them. A box containing my pipetmen and a few labeled tubes was ceremoniously deposited in my office, without explanation; the message was clear!) It was in this paper that we documented the first of the chromosome-specific alpha satellite DNAs and proposed that "a collection of repeated DNA probes specific for each human chromosome might be useful for molecular cytogenetic analysis in certain clinical situations."[34] But with autoradiography and radioactive in situ hybridization (Figure 7A), this approach was never going to enjoy widespread usage in clinical situations. It took Dan Pinkel and Joe Gray, who were wonderfully welcoming as collaborators, to help us realize that these probes could be used as fluorescence markers for individual chromosomes[35] (Figure 7B). It was Dan and Joe's efforts, of course, that led to the complete revolution in clinical cytogenetic analysis at that time.[36]

So just as Carolyn Brown was the critical individual who transformed the study of X inactivation, it was John Waye who was the critical individual for opening the study of centromeres and alpha satellite during the mid-to-late 1980s. In that pre-computational era—recall that no one had heard of computational biology in those days; sequence analysis to us was having scrolls of DNA

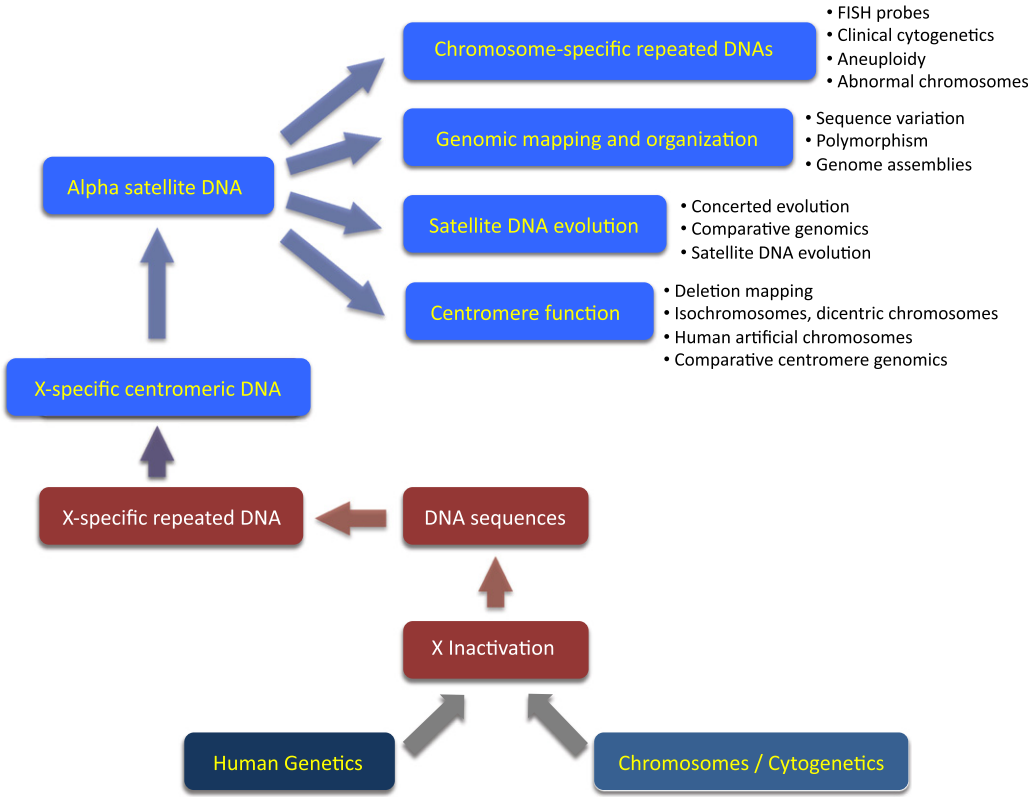

Figure 6. Intellectual Flow Diagram: Centromeres

sequence rolled out across the living room floor and a pencil and a piece of paper, trying to figure out how it all worked—it was John who recognized the hierarchical organization of alpha satellite and who published a series of some 16 papers as a graduate student at the University of Toronto. He established a conceptual framework[37–39] for the study of satellite DNAs in the human genome that is still guiding the field today, and, like all good graduate students, he brought in the next group of students, Sharon Durfy and Melanie Mahtani, to pursue their own questions about alpha satellite.[40,41]

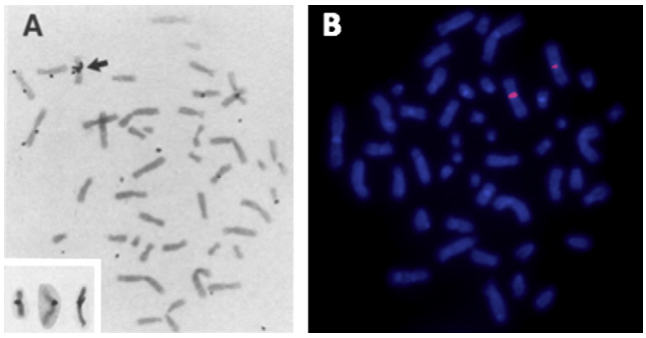

Figure 7. Chromosome-Specific Alpha Satellite DNA in the Human Genome
(A) In situ hybridization of ^{3}H-labeled alpha satellite from the X chromosome to metaphase chromosomes from a male cell line. Arrow indicates the X chromosome. Inset shows X chromosomes from additional cells. Reprinted from [34].
(B) Fluorescence in situ hybridization of alpha satellite from chromosome 3 to metaphase chromosomes from a male cell line. (Image courtesy of K. Hayden, Duke University.)

While the genomics, genetics, and molecular cytogenetics of alpha satellite had its foundations in evolutionary thought, the real work on the evolution of these sequences didn't begin until Peter Warburton joined the lab in Toronto. For Peter, this was like joining the family business, having inherited a love of chromosomes from his mother and an understanding of evolutionary biology from his father. He took ownership of a series of projects, not so much because he desperately wanted to understand evolution by itself, but because he had the insight to realize that understanding the evolution of these sequences would be critical to finally getting us to what he called the "FMC," the functional mammalian centromere. He turned out to be right, and that intellectual connection still drives much of the thinking on centromeric DNAs today in a host of organisms. His early studies to identify exactly how these sequences were organized across the many megabases of alpha satellite at our centromeric regions[42] and how the organization of those sequences might relate to centromere function[43] reflect questions that we still don't know the answers to, although, now some 20 years later, the ever-improving tools of genomic analysis and DNA sequencing may finally allow us to test some of the hypotheses put forward back in the late 1980s (Figure 8).

Notwithstanding the importance of these early studies for understanding centromeric DNA, none of it had to do

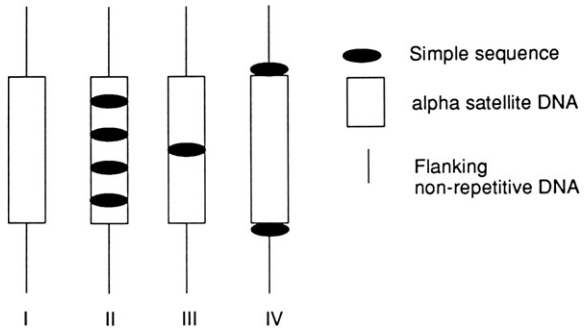

Figure 8. Models for the Organization of DNA at Human Centromeres

Four simple models for the structure of the human centromere are presented, based on [45]. In 1989, alpha satellite was hypothesized to occupy the central position critical to centromere function, and four models were presented to guide future genomic and functional analyses. In model I, alpha satellite *is* the functional centromere and is the only sequence required for its activity. Models II, III, and IV posit additional functional elements (present in one, two, or multiple copies). Current cytogenetic and immunofluorescence data, chromatin immunoprecipitation experiments, deletion mapping, and human artificial chromosome assays all support model I. (Figure from Ph.D. thesis of R. Wevrick, University of Toronto, 1992.)

with centromere function directly. It was Rachel Wevrick who joined the lab and was the first one to focus not solely on genomics and evolution but on functional aspects of the story. She carried out a wonderfully thorough study of a chromosome abnormality that Patricia Howard-Peebles had identified to get us to understand exactly happens when a chromosome break occurs in the middle of a satellite array and in the middle of the centromere.[44] This rearrangement showed us that, just like an earth worm, if you took a centromere and split it in half, you would end up with two equally functional centromeres: in this case, one a marker chromosome and the other one with a deletion of chromosome 17, both of which were functional. This was a wonderful illustration of the power of human genetic material and the first formal demonstration that complex centromeres in our and other complex genomes were functionally repetitive, not just structurally repetitive.

This finding led to the hypothesis that Rachel, Peter and I finally put forward publicly in 1989: that in fact it was the satellite DNA sequences themselves that were responsible for centromere function in complex genomes,[45] an idea that would have been anathema to the field just a few years earlier. This hypothesis led to a variety of studies in the lab to begin to investigate aspects of centromere function and to test predictions of some of the models we put forward (Figure 8). Thomas Haaf and Peter showed that alpha satellite sequences could behave as centromere sequences when introduced into cells in culture.[46] This provided an impetus for John Harrington and Gil van Bokkelen to generate the first human artificial chromosomes in 1997 and thus demonstrate that alpha satellite alone was capable of providing all the genomic instructions needed to form a functional centromere.[47]

This was by no means the entire story, however, and a whole series of studies of the epigenetic and genomic aspects of centromeres were carried out by Beth Sullivan, by Mary Schueler, by Anne Higgins, and by Katie Rudd to finally provide the genetic and genomic evidence needed to say that alpha satellite was in fact the functional centromere in normal human chromosomes.[48–51] I hasten to add, however, that, among others, it is Beth Sullivan and Peter Warburton who continue to challenge the apparent simplicity of this conclusion by pointing to the very clear role of epigenetics in specifying centromere and neocentromere function in various abnormal human chromosomes. There's plenty of unfinished business here!

From a genomic perspective, it was Mary and Katie who recognized that, in order to finally pull this story together, we were going to have to tackle the part of the Human Genome Project that was left behind, the large gaps in the middle of chromosomes that were never part of the original genome project.[17,52,53] They have since passed the ball on to the next generation of students, still in The Sandbox now, who are working feverishly to take now complete sequences of the human genome and begin to fill in those centromere gaps, so we can finally understand the identity and organization of sequences that make up our centromeres.

From Genomes to Biology: The Chromosome as Integrator

If there have been two principal areas of study in The Sandbox up to now, there remains a third bit of unfinished business. How we go from primary genome sequence to get to biology, passing through the many states of chromatin that are the focus of much current attention, remains still very much in the realm of "and then a miracle occurs." As I view this opportunity for synthesis, the chromosome is still the center of action and the source of integration, both physically and functionally. From my admittedly somewhat parochial perspective as a "chromosome person," we have yet to fully understand the nature of centromeres and the nature of their epigenetic modifications; how heterochromatin forms, spreads, and gets maintained; and how heterochromatin can lead to gene silencing in steps that almost certainly involve noncoding RNAs. Yet, we also have to account for regions of the genome that somehow manage to avoid the insidious effect of epigenetic silencing. What is it about those sequences or about the organization or folding of a genome and of a chromosome that allows those genes to continue to be expressed despite living in an environment that is otherwise quite inhospitable to gene function?

To me, it is this overall synthesis that remains the most challenging experiment yet to be done, trying to understand how we go from chromosomes and genome sequences to the underlying biology of the cell. After some 30-plus years, it is this integration that remains our most pressing unfinished business.

Acknowledgments and Thanks

I'm enormously grateful to a number of colleagues and friends over the years, many of whom are here in this room today—those from the very early days when I started The Sandbox in Toronto, those who have shared my enthusiasm for sex chromosome biology and evolution, those who have helped us think about chromosomes and genomes over the years, those who have served as exemplars of how to lead a life in science, and those who serve as ongoing mentors, most especially my once and always mentor, Lee Rosenberg.

I want to acknowledge in particular the group of individuals who joined me in Cleveland during our years there and remind me regularly of what was so special about those years for all of us. Even though many have since scattered to other institutions, l will say this about those years—and I'm willing to guess that I'm the first person to ever say this line in public— "we will always have Cleveland!"

This story has been, as I said at the outset, the story of a number of different students, and I've tried to highlight many of them who have been so critical for defining our paths on these projects. I'm so grateful to those groups of individuals, students and others, who started with us in Toronto, moved to Stanford for a few years before earthquakes chased us back east, and then moved to Cleveland to Case Western Reserve University. As well, of course, I'm grateful to the current members of The Sandbox at Duke, who have both benefitted from and remain challenged by the work of those who were in the lab before them. Ultimately, I am sure that it is the members of The Sandbox who will mean so much more to me and to the field than whatever textbooks may or may not say about X inactivation or the *XIST* gene or centromeres in human and other chromosomes. This is the group that is responsible for this story and the group who will pass it down to their students, as they pursue answers to their own questions.

And lastly, one more person to thank: the most important person in this whole story has been Vicki Powers, my wife of now some 30 years without whom none of this would be either worth it or possible. We met in the laboratory and worked side-by-side in The Sandbox for many of those years. Eight years ago, I stood in front of you as your President and said that you represented my community, as you do. But if you are my community, it is she who has been my world, and I am more grateful to her than she can possibly know.

You all have made me enormously proud today. I thank you for this honor, and I thank you so very much for honoring the people who've worked with me all these years.

References

1. Stern, C. (1973). Principles of Human Genetics, Third Edition (San Francisco: W.H. Freeman and Company), p. 872.
2. Lyon, M.F. (1961). Gene action in the X-chromosome of the mouse (Mus musculus L.). Nature *190*, 372–373.
3. Brown, S.W., and Chandra, H.S. (1973). Inactivation system of the mammalian X chromosome. Proc. Natl. Acad. Sci. USA *70*, 195–199.
4. Therman, E., and Patau, K. (1974). Abnormal X chromosomes in man: origin, behavior and effects. Humangenetik *25*, 1–16.
5. Therman, E., Sarto, G.E., and Patau, K. (1974). Center for Barr body condensation on the proximal part of the human Xq: a hypothesis. Chromosoma *44*, 361–366.
6. Latt, S.A. (1973). Microfluorometric detection of deoxyribonucleic acid replication in human metaphase chromosomes. Proc. Natl. Acad. Sci. USA *70*, 3395–3399.
7. Willard, H.F. (1989). In memory of Samuel A. Latt, M.D. Ph.D. Am. J. Hum. Genet. *44*, 766–767.
8. Willard, H.F., and Latt, S.A. (1976). Analysis of DNA replication in human X chromosomes by fluorescence microscopy. Am. J. Hum. Genet. *28*, 213–227.
9. Willard, H.F. (2002). 2001 ASHG Presidential Address. On black boxes and storytellers: lessons learned in human genetics. Am. J. Hum. Genet. *70*, 285–296.
10. Willard, H.F. (1977). Tissue-specific heterogeneity in DNA replication patterns of human X chromosomes. Chromosoma *61*, 61–73.
11. Willard, H.F., and Breg, W.R. (1980). Human X chromosomes: synchrony of DNA replication in diploid and triploid fibroblasts with multiple active or inactive X chromosomes. Somat. Cell Genet. *6*, 187–198.
12. Brown, C.J., and Willard, H.F. (1990). Localization of a gene that escapes inactivation to the X chromosome proximal short arm: implications for X inactivation. Am. J. Hum. Genet. *45*, 273–279.
13. Brown, C.J., Carrel, L., and Willard, H.F. (1997). Expression of genes from the human active and inactive X chromosomes. Am. J. Hum. Genet. *60*, 1333–1343.
14. Carrel, L., Cottle, A.A., Goglin, K.C., and Willard, H.F. (1999). A first-generation X-inactivation profile of the human X chromosome. Proc. Natl. Acad. Sci. USA *96*, 14440–14444.
15. Miller, A.P., and Willard, H.F. (1998). Chromosomal basis of X chromosome inactivation: identification of a multigene domain in Xp11.21-p11.22 that escapes X inactivation. Proc. Natl. Acad. Sci. USA *95*, 8709–8714.
16. Carrel, L., and Willard, H.F. (2005). X-inactivation profile reveals extensive variability in X-linked gene expression in females. Nature *434*, 400–404.
17. Ross, M.T., Grafham, D.V., Coffey, A.J., Scherer, S., McLay, K., Muzny, D., Platzer, M., Howell, G.R., Burrows, C., Bird, C.P., et al. (2005). The DNA sequence of the human X chromosome. Nature *434*, 325–337.
18. Lyon, M.F. (1964). Second International Conference on Congenital Malformations, Session I, Discussion, pp. 67–68. New York: International Medical Congress Ltd.
19. Brown, C.J., Lafreniere, R.G., Powers, V.E., Sebastio, G., Ballabio, A., Pettigrew, A.L., Ledbetter, D.H., Levy, E., Craig, I.W., and Willard, H.F. (1991). Localization of the X inactivation centre on the human X chromosome in Xq13. Nature *349*, 82–84.
20. Brown, C.J., Ballabio, A., Rupert, J.L., Lafreniere, R.G., Grompe, M., Tonlorenzi, R., and Willard, H.F. (1991). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature *349*, 38–44.

21. Brown, C.J., Hendrich, B.D., Rupert, J.L., Lafrenière, R.G., Xing, Y., Lawrence, J., and Willard, H.F. (1992). The human *XIST* gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. Cell *71*, 527–542.
22. Brannan, C.I., Dees, E.C., Ingram, R.S., and Tilghman, S.M. (1990). The product of the H19 gene may function as an RNA. Mol. Cell. Biol. *10*, 28–36.
23. Plenge, R.M., Hendrich, B.D., Schwartz, C., Arena, J.F., Naumova, A., Sapienza, C., Winter, R.M., and Willard, H.F. (1997). A promoter mutation in the *XIST* gene in two unrelated families with skewed X-chromosome inactivation. Nat. Genet. *17*, 353–356.
24. Amos-Landgraf, J.M., Cottle, A., Plenge, R.M., Friez, M., Schwartz, C.E., Longshore, J., and Willard, H.F. (2006). X chromosome inactivation patterns of 1005 phenotypically unaffected females. Am. J. Hum. Genet. *79*, 493–499.
25. Plenge, R.M., Stevenson, R.A., Lubs, H.A., Schwartz, C.E., and Willard, H.F. (2002). Skewed X-chromosome inactivation is a common feature of X-linked mental retardation disorders. Am. J. Hum. Genet. *71*, 168–173.
26. Chadwick, B.P., and Willard, H.F. (2003). Chromatin of the Barr body: histone and non-histone proteins associated with or excluded from the inactive X chromosome. Hum. Mol. Genet. *12*, 2167–2178.
27. Chadwick, B.P., and Willard, H.F. (2004). Multiple spatially distinct types of facultative heterochromatin on the human inactive X chromosome. Proc. Natl. Acad. Sci. USA *101*, 17450–17455.
28. Valley, C.M., Pertz, L.M., Balakumaran, B.S., and Willard, H.F. (2006). Chromosome-wide, allele-specific analysis of the histone code on the human X chromosome. Hum. Mol. Genet. *15*, 2335–2347.
29. Kunkel, L.M., Smith, K.D., and Boyer, S.H. (1976). Human Y-chromosome-specific reiterated DNA. Science *191*, 1189–1190.
30. Schmeckpeper, B.J., Willard, H.F., and Smith, K.D. (1981). Isolation and characterization of cloned human DNA fragments carrying reiterated sequences common to both autosomes and the X chromosome. Nucleic Acids Res. *9*, 1853–1872.
31. Gartler, S.M., and Riggs, A.D. (1983). Mammalian X-chromosome inactivation. Annu. Rev. Genet. *17*, 155–190.
32. Willard, H.F., Smith, K.D., and Sutherland, J. (1983). Isolation and characterization of a major tandem repeat family from the human X chromosome. Nucleic Acids Res. *11*, 2017–2033.
33. Waye, J.S., and Willard, H.F. (1985). Chromosome-specific alpha satellite DNA: Nucleotide sequence analysis of the 2.0 kilobasepair repeat from the human X chromosome. Nucleic Acids Res. *12*, 2731–2743.
34. Willard, H.F. (1985). Chromosome-specific organization of human alpha satellite DNA. Am. J. Hum. Genet. *37*, 524–532.
35. Waye, J.S., Durfy, S.J., Pinkel, D., Kenwrick, S., Patterson, M., Davies, K.E., and Willard, H.F. (1987). Chromosome-specific alpha satellite DNA from human chromosome 1: hierarchical structure and genomic organization of a polymorphic domain spanning several hundred kilobase pairs of centromeric DNA. Genomics *1*, 43–51.
36. Pinkel, D., Straume, T., and Gray, J.W. (1986). Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. Proc. Natl. Acad. Sci. USA *83*, 2934–2938.
37. Waye, J.S., and Willard, H.F. (1986). Structure, organization, and sequence of alpha satellite DNA from human chromosome 17: evidence for evolution by unequal crossing-over and an ancestral pentamer repeat shared with the human X chromosome. Mol. Cell. Biol. *6*, 3156–3165.
38. Willard, H.F., and Waye, J.S. (1987). Hierarchical order in chromosome-specific human alpha satellite DNA. Trends Genet. *3*, 192–198.
39. Waye, J.S., and Willard, H.F. (1989). Human beta satellite DNA: genomic organization and sequence definition of a class of highly repetitive tandem DNA. Proc. Natl. Acad. Sci. USA *86*, 6250–6254.
40. Durfy, S.J., and Willard, H.F. (1989). Patterns of intra- and interarray sequence variation in alpha satellite from the human X chromosome: evidence for short-range homogenization of tandemly repeated DNA sequences. Genomics *5*, 810–821.
41. Mahtani, M.M., and Willard, H.F. (1990). Pulsed-field gel analysis of alpha-satellite DNA at the human X chromosome centromere: high-frequency polymorphisms and array size estimate. Genomics *7*, 607–613.
42. Warburton, P.E., and Willard, H.F. (1990). Genomic analysis of sequence variation in tandemly repeated DNA. Evidence for localized homogeneous sequence domains within arrays of alpha-satellite DNA. J. Mol. Biol. *216*, 3–16.
43. Warburton, P.E., Waye, J.S., and Willard, H.F. (1993). Nonrandom localization of recombination events in human alpha satellite repeat unit variants: implications for higher-order structural characteristics within centromeric heterochromatin. Mol. Cell. Biol. *13*, 6520–6529.
44. Wevrick, R., Earnshaw, W.C., Howard-Peebles, P.N., and Willard, H.F. (1990). Partial deletion of alpha satellite DNA associated with reduced amounts of the centromere protein CENP-B in a mitotically stable human chromosome rearrangement. Mol. Cell. Biol. *10*, 6374–6380.
45. Willard, H.F., Wevrick, R., and Warburton, P.E. (1989). Human centromere structure: organization and potential role of alpha satellite DNA. In Aneuploidy: Mechanisms of Origin, M.A. Resnick, ed. (New York: Liss, Inc.), pp. 9–18.
46. Haaf, T., Warburton, P.E., and Willard, H.F. (1992). Integration of human alpha-satellite DNA into simian chromosomes: centromere protein binding and disruption of normal chromosome segregation. Cell *70*, 681–696.
47. Harrington, J.J., Van Bokkelen, G., Mays, R.W., Gustashaw, K., and Willard, H.F. (1997). Formation of *de novo* centromeres and construction of first-generation human artificial microchromosomes. Nat. Genet. *15*, 345–355.
48. Sullivan, B.A., and Willard, H.F. (1998). Stable dicentric X chromosomes with two functional centromeres. Nat. Genet. *20*, 227–228.
49. Schueler, M.G., Higgins, A.W., Rudd, M.K., Gustashaw, K., and Willard, H.F. (2001). Genomic and genetic definition of a functional human centromere. Science *294*, 109–115.
50. Higgins, A.W., Gustashaw, K.M., and Willard, H.F. (2005). Engineered human dicentric chromosomes show centromere plasticity. Chromosome Res. *13*, 745–762.
51. Rudd, M.K., Mays, R.W., Schwartz, S., and Willard, H.F. (2003). Human artificial chromosomes with alpha satellite-based *de novo* centromeres show increased frequency of nondisjunction and anaphase lag. Mol. Cell. Biol. *23*, 7689–7697.
52. Rudd, M.K., Schueler, M.G., and Willard, H.F. (2004). Characterization and functional annotation of human centromeres. Cold Spring Harb. Symp. Quant. Biol. *68*, 141–150.
53. Rudd, M.K., and Willard, H.F. (2004). Analysis of the centromeric regions of the human genome assembly. Trends Genet. *20*, 529–533.