Tandem Repeat Polymorphisms

ADVANCES IN EXPERIMENTAL MEDICINE AND BIOLOGY

Recent Volumes in this Series

Volume 760
REGENERATIVE BIOLOGY OF THE SPINE AND SPINAL CORD
Rahul Jandial and Mike Y. Chen

Volume 761
OOCYTE BIOLOGY AND FERTILITY PRESERVATION
Sam Kim

Volume 762
HIV INTERACTIONS WITH DENDRITIC CELLS: INFECTION AND IMMUNITY
Li Wu and Olivier Schwartz

Volume 763
BIOLOGY AND REGULATION OF BLOOD-TISSUE BARRIERS
C. Yan Cheng

Volume 764
HOT TOPICS IN INFECTION AND IMMUNITY IX
Andrew J. Pollard

Volume 765
OXYGEN TRANSPORT TO TISSUE XXXIV
Palm Welch and Harrison Bruley

Volume 766
DIGENETIC TREMATODES
Rafael Toledo and Bernard Fried

Volume 767
DNA HELICASES AND DNA MOTOR PROTEINS
Maria Spies

Volume 768
TEN YEARS OF PROGRESS IN GW/P BODY RESEARCH
Edward K.L. Chan and Marvin J. Fritzler

Volume 769
TANDEM REPEAT POLYMORPHISMS: GENETIC PLASTICITY, NEURAL DIVERSITY AND DISEASE
Anthony J. Hannan

Tandem Repeat Polymorphisms

Genetic Plasticity, Neural Diversity and Disease

Edited by

Anthony J. Hannan, PhD
Florey Neuroscience Institutes, Melbourne Brain Centre, University of Melbourne, Parkville, Victoria, Australia

Springer Science+Business Media, LLC

Landes Bioscience

Springer Science+Business Media, LLC
Landes Bioscience

Springer Science+Business Media, LLC, 233 Spring Street, New York, New York 10013, USA
http://www.springer.com

Please address all inquiries to the publishers:
Landes Bioscience, 1806 Rio Grande, Austin, Texas 78701, USA
Phone: 512/ 637 6050; FAX: 512/ 637 6079
http://www.landesbioscience.com

The chapters in this book are available in the Madame Curie Bioscience Database.
http://www.landesbioscience.com/curie

Tandem Repeat Polymorphisms: Genetic Plasticity, Neural Diversity and Disease, edited by Anthony J. Hannan. Landes Bioscience / Springer Science+Business Media, LLC dual imprint / Springer series: Advances in Experimental Medicine and Biology.

ISBN 978-1-4614-5433-5

Library of Congress Cataloging-in-Publication Data

Tandem repeat polymorphisms : genetic plasticity, neural diversity, and disease / edited by Anthony J. Hannan.
p. ; cm. -- (Advances in experimental medicine and biology ; v. 769)
Includes bibliographical references and index.
ISBN 978-1-4614-5433-5 (alk. paper)
I. Hannan, Anthony J., 1969- II. Series: Advances in experimental medicine and biology ; v. 769. 0065-2598
[DNLM: 1. Polymorphism, Genetic. 2. Tandem Repeat Sequences. 3. Heredodegenerative Disorders, Nervous System--genetics. 4. Mutation. W1 AD559 v.769 2012 / QU 500]

576.5’49--dc23

2012030285

DEDICATION

To all of those families who are suffering from tandem repeat expansion disorders

PREFACE

Tandem repeats of DNA sequences provide a unique and abundant source of genomic variability and recent evidence suggests they can modulate a range of biological processes in a wide variety of different species. These classes of repetitive DNA are variously referred to as simple sequence repeats, satellite DNA (microsatellites, minisatellites and satellites) or variable number tandem repeats. A key aspect of tandem repeats is that they represent highly polymorphic and uniquely mutable genomic components which can (depending on their sequence, length and location) affect the structure and function of DNA, RNA and protein.

This book addresses the role of tandem repeat polymorphisms (TRPs) in genetic plasticity, evolution, development, biological processes, neural diversity, brain function, dysfunction and disease. There are hundreds of thousands of unique tandem repeats in the human genome and their polymorphic distributions have the potential to greatly influence functional diversity and disease susceptibility. Recent discoveries in this expanding field are critically reviewed and discussed in a range of subsequent chapters, with a focus on the role of TRPs and their various gene products in evolution, development, diverse molecular and cellular processes, brain function and disease.

In the first chapter, I introduce these broad themes. This includes discussion of the specific proposal that TRPs could help solve the conundrum of 'missing heritability' produced by genome-wide association studies of various polygenic complex diseases which have only examined single nucleotide polymorphisms (SNPs). Subsequent chapters focus on key aspects of TRPs in health and disease. In the second chapter, David King shares his ideas regarding the role of simple sequence repeats in evolution, and provides a detailed discussion of repeat sequences as mutable sites providing genetic variability upon which natural selection can act. This theme is then extended by Noel Faux, who uses bioinformatic analyses of trinucleotide repeats encoding homopeptides to explore both the evolution and function of a wide variety of amino acid repeats located in diverse proteins across the phylogenetic spectrum. This bioinformatic exploration of the role of TRPs in normal biological functions is then extended by Sterling Sawaya and colleagues, who discuss evidence for the role of promoter microsatellites in modulating the expression of various human genes.

Expansions in tandem repeats ('dynamic mutations') are known to cause many disorders, which mainly affect the nervous system, including Huntington's disease (the most common polyglutamine disorder), spinocerebellar ataxias, Kennedy's disease (spinobulbar muscular atrophy), dentatorubral-pallidoluysian atrophy, Friedreich ataxia, polyalanine disorders, fragile X syndrome and related disorders. Robert Richards, who helped coin the term dynamic mutations, and his colleague Clare van Eyk, have provided an overview of this large and clinically significant area of research. Whilst the plasticity of these tandem repeats occurs at the DNA level, evidence for both 'gain of function' and 'loss of function' pathogenic effects of repeat expansions ('genetic stutters') at RNA and protein levels is discussed using specific examples of these monogenic disorders.

Danuta Loesch and Randi Hagerman review the exciting field that has evolved around the FMR1 gene, originally discovered due to its hosting of the large expansion of a 5'UTR trinucleotide (CGG) repeat which causes fragile X syndrome. Smaller 'premutation' repeat lengths have recently been shown by Hagerman and colleagues to cause fragile X tremor-ataxia syndrome (FXTAS), as well as contributing to other disorders. The focus then shifts to protein-coding trinucleotide repeats, with complementary chapters from Amy Robertson and Stephen Bottomley, as well as Saski Polling, Andrew Hill and Danny Hatters, exploring the biochemistry of expanded polyglutamine tracts and their roles in at least nine autosomal dominant neurodegenerative disorders. As Huntington's disease (HD), which was first described by George Huntington in 1872 and genetically mapped over a century later, is the most common of these so-called polyglutamine disorders it has been most intensively researched. Henry Waldvogel and colleagues review data on the neuropathology and related symptomatology of HD, linking molecular and cellular aspects to pathogenesis at the systems level. Another extraordinary polyglutamine disorder is Kennedy's disease (also known as spinobulbar muscular atrophy or SBMA). Jeffrey Zajac and Mark Tang discuss how polyglutamine polymorphism in the androgen receptor not only causes SBMA, but can contribute to other complex disorders via modulation of this sex hormone signaling system. A different neurodegenerative disease caused by a non-coding tandem repeat expansion is Friedreich ataxia. Corben and colleagues discuss how homozygosity of a GAA repeat expansion in an intron of the frataxin gene leads to downregulated expression and consequent neuropathology, motor and cognitive symptoms. In the final chapter, a unique group of disorders involving trinucleotide repeat expansions encoding polyalanine tracts, are reviewed by Cheryl Shoubridge and Jozef Gecz, providing insights into how expanded polyalanine in specific proteins leads to developmental abnormalities and neurocognitive dysfunction.

It is hoped that this book will help the reader to grasp the significance of TRPs in evolution, development, brain function and a variety of major clinical disorders. As we begin a new genetic revolution powered by next-generation sequencing, expanding knowledge of tandem repeats and their polymorphic variants will no doubt continue to enhance our understanding of genetic plasticity, neural diversity and disease.

Anthony J. Hannan, PhD
Florey Neuroscience Institute,
Melbourne Brain Centre
University of Melbourne
Parkville, Victoria, Australia

ABOUT THE EDITOR...

ANTHONY J. HANNAN is Head of the Neural Plasticity Laboratory, Florey Neuroscience Institutes and Associate Professor at the University of Melbourne, Parkville, Australia. Following undergraduate and PhD degrees at the University of Sydney, Anthony received postdoctoral neuroscience training at the University of Oxford, supported by a Nuffield Medical Fellowship. He currently holds an Australian Research Council (ARC) Future Fellowship (FT3) and an Honorary Senior Research Fellowship from the National Health and Medical Research Council (NHMRC). Main research interests include pathogenic mechanisms mediating Huntington's disease and related tandem repeat expansion disorders, as well as other cognitive and psychiatric illnesses. In his laboratory, experimental models of gene-environment interactions are used to explore experience-dependent plasticity in the healthy and diseased brain.

PARTICIPANTS

Andrew T. Bagshaw
Department of Pathology
University of Otago-Christchurch
Christchurch
New Zealand

Stephen P. Bottomley
Department of Biochemistry
and Molecular Biology
Monash University
Clayton, Victoria
Australia

John L. Bradshaw
Experimental Neuropsychology
Research Unit
School of Psychology and Psychiatry
Monash University
Clayton, Victoria
Australia

Emmanuel Buschiazzo
School of Natural Sciences
University of California – Merced
Merced, California
USA

Andrew J. Churchyard
Monash Neurology
Monash Medical Centre
Clayton, Victoria
Australia

Louise A. Corben
Bruce Lefroy Centre for Genetic
Health Research
Murdoch Childrens Research Institute
The Royal Children's Hospital
Parkville, Victoria
Australia

Martin B. Delatycki
Department of Clinical Genetics
Austin Health
and
Department of Medicine
University of Melbourne at Austin Health
Heidelberg, Victoria
Australia

Marguerite V. Evans-Galea
Bruce Lefroy Centre for Genetic
Health Research
Murdoch Childrens Research Institute
The Royal Children's Hospital
Parkville, Victoria
Australia

Richard L.M. Faull
Department of Anatomy with Radiology
Faculty of Medical and Health Sciences
and
Centre for Brain Research
University of Auckland
Auckland
New Zealand

Noel Faux
Mental Health Research Institute
The University of Melbourne
Parkville, Victoria
and
National Neuroscience Facility
Carlton, Victoria
Australia

Mark Ng Tang Fui
Department of Endocrinology
Austin Health
Heidelberg, Victoria
Australia

Jozef Gecz
Department of Genetics
and Molecular Pathology
SA Pathology at the Women's
and Children's Hospital
North Adelaide, South Australia
and
Department of Pediatrics
University of Adelaide
Adelaide, South Australia
Australia

Neil J. Gemmell
Department of Anatomy
and Structural Biology
University of Otago
Dunedin
New Zealand

Nellie Georgiou-Karistianis
Experimental Neuropsychology
Research Unit
School of Psychology and Psychiatry
Monash University
Clayton, Victoria
Australia

Randi Hagerman
Department of Pediatrics
MIND Institute
University of California at Davis
Medical Center
Sacramento, California
USA

Anthony J. Hannan
Florey Neuroscience Institutes
Melbourne Brain Centre
University of Melbourne
Parkville, Victoria
Australia

Danny M. Hatters
Department of Biochemistry
and Molecular Biology
University of Melbourne
Melbourne, Victoria
Australia

Andrew F. Hill
Department of Biochemistry
and Molecular Biology
University of Melbourne
Melbourne, Victoria
Australia

Virginia Hogg
Lynette Tippett
Centre for Brain Research
and
Department of Psychology
University of Auckland
Auckland
New Zealand

David G. King
Department of Anatomy
and
Department of Zoology
Southern Illinois University Carbondale
Carbondale, Illinois
USA

Danuta Loesch
Department of Psychology
LaTrobe University
Melbourne, Victoria
Australia

Saskia Polling
Department of Biochemistry
and Molecular Biology
University of Melbourne
Melbourne, Victoria
Australia

Robert I. Richards
Discipline of Genetics
School of Molecular
and Biomedical Sciences
The University of Adelaide
Adelaide, South Australia
Australia

Amy L. Robertson
Department of Biochemistry
and Molecular Biology
Monash University
Clayton, Victoria
Australia

Sterling M. Sawaya
Department of Anatomy
and Structural Biology
University of Otago
Dunedin
New Zealand

Cheryl Shoubridge
Department of Genetics
and Molecular Pathology
SA Pathology at the Women's
and Children's Hospital
North Adelaide, South Australia
and
Department of Pediatrics
University of Adelaide
Adelaide, South Australia
Australia

Doris Thu
Brain Mind Institute
Ecole Polytechnique Federale de Lausanne
Lausanne
Switzerland

Lynette Tippett
Centre for Brain Research
and
Department of Psychology
University of Auckland
Auckland
New Zealand

Clare L. van Eyk
Discipline of Genetics
School of Molecular
and Biomedical Sciences
The University of Adelaide
Adelaide, South Australia
Australia

Henry J. Waldvogel
Department of Anatomy with Radiology
Faculty of Medical and Health Sciences
and
Centre for Brain Research
University of Auckland
Auckland
New Zealand

Jeffrey D. Zajac
Department of Medicine
and
Department of Endocrinology
University of Melbourne at Austin Health
Heidelberg, Victoria
Australia

CONTENTS

ACKNOWLEDGMENTS

I would firstly like to thank the authors who have written chapters for this book, dedicating substantial amounts of precious time in their already hectic schedules. I would like to thank the dedicated staff at Landes Bioscience, particularly Celeste Carlton and Cynthia Conomos, for their assistance and patience during this long journey. I greatly appreciate ongoing discussions with past and present members of my laboratory, as well as other colleagues, at the Florey Neuroscience Institutes, University of Melbourne, as the intellectual stimulation they have provided has contributed to the development of my own thoughts in the introductory chapter. I thank Carolyn M. Hannan for her support with editing. My thanks also to the funding bodies that have allowed me to devote so much of my time to research, in particular the Australian Research Council and the National Health and Medical Research Council.

CHAPTER 1

TANDEM REPEAT POLYMORPHISMS

Mediators of Genetic Plasticity, Modulators of Biological Diversity and Dynamic Sources of Disease Susceptibility

Anthony J. Hannan
Florey Neuroscience Institutes, Melbourne Brain Centre, University of Melbourne, Parkville, Australia
Email: anthony.hannan@florey.edu.au

Abstract: Tandem repetitive DNA elements (tandem repeats), including microsatellites and simple sequence repeats, are extremely common throughout the genomes of a wide range of species. Tandem repeat expansions have been found to cause a range of monogenic diseases, such as Huntington's disease, various ataxias and other neurological diseases. The human genome contains hundreds of thousands of distinct tandem repeats, many of which appear to have evolved to regulate specific aspects of gene expression, RNA function and protein function. Tandem repeat polymorphisms (TRPs) provide a unique source of genetic variability that has an extended digital distribution, as opposed to the usual binary nature of single nucleotide polymorphisms. In this chapter I will review studies in which tandem repeats have been implicated in a multitude of molecular and cellular processes associated with the development, behavior and evolution of a variety of animal species, including mammals. Recent data suggesting that these repetitive sequences can increase the 'evolvability' of genomes provides further evidence that TRPs not only have functional consequences but also provide a rich source of genetic diversity that can facilitate evolutionary processes. I propose that a readily mutable subclass of tandem repeats may provide an important template for stochastic genetic variation, which could in turn generate diversity in epigenetics, development and organismal function, thus impacting upon evolution. Furthermore, the distinctive characteristics of TRPs also uniquely position them as contributors to complex polygenic disorders. Ultimately, there is much to be gained from systematic analysis of the 'repeatome', defined as the entire set of tandem repeats and other repetitive DNA in a genome, as well as their transcribed and translated expression products. Applying such approaches not only to the human genome but to other species will yield new insights into the genetic regulation of a wide range of biological processes in healthy and diseased states.

Tandem Repeat Polymorphisms: Genetic Plasticity, Neural Diversity and Disease,
edited by Anthony J. Hannan ©2012 Landes Bioscience and Springer Science+Business Media.

INTRODUCTION

DNA sequencing technologies have led to a recent revolution in our understanding of the human genome, as well as the genomes of other animals and many other species. Tandemly repeated DNA sequences, or tandem repeats, are increasingly being recognized as much more than 'genetic stutters', but rather key structural and functional elements of the human genome, as well as the genomes of other species.[1-4] The term tandem repeats encompasses satellite DNA (including minisatellites and microsatellites), simple sequence repeats (SSRs), as well as variable number tandem repeats (VNTRs). Tandem repeats are found commonly in exons, as well as introns and intergenic regions. Whilst the function of tandem repeats has only been explored in a relatively small number of genes, their abundance and locations suggest potentially widespread roles in the modulation of gene expression, RNA function, protein function and other molecular processes.[1,5-10] Furthermore, tandem repeats located in transcribed regions of the genome, which may constitute a large proportion of the human genome[11] have the capacity to alter the structure and function of both coding and noncoding RNA species. Those tandem repeats located in coding regions (and encoding amino acid repeats via trinucleotides, hexanucleotides, etc.) have additional potential roles in modulating the structure and function of the encoded proteins.

The importance of tandem repeats can be recognized in a number of different domains. The high degree of polymorphism in tandem repeats may confer a capacity to act as 'tuning knobs' for evolutionary processes.[1-6,12] This capacity derives from the fact that tandem repeats are far more mutable than single nucleotides, meaning that tandem repeat polymorphisms (TRPs) can have an extended digital distribution[2] as opposed to the binary possibilities presented by individual single nucleotide polymorphisms (SNPs). The mutability of tandem repeats has been proposed to add variability to brain development and function, thus providing a more dynamic template upon which natural selection may act.[1]

The capacity to compare the genomes of different individuals of a species has not only revealed high levels of conservation of many genes and intergenic regions, but also the extent of polymorphism.[13] While much attention has been focused on SNPs, recent studies have revealed other important polymorphic DNA sequences, including TRPs[2] and copy number variants (CNVs).[14]

EXPANSION OF TANDEM REPEATS IN RARE MONOGENIC DISORDERS WITH MENDELIAN INHERITANCE

The term 'dynamic mutations' has been used to describe the expansions of tandem repeats associated with a variety of diseases exhibiting Mendelian inheritance patterns.[15] TRPs, above a specific repeat length, have been shown to cause various autosomal dominant and recessive human disorders, including polyglutamine disorders (e.g., Huntington's disease and some spinocerebellar ataxias), Friedreich ataxia, fragile X syndrome and myotonic dystrophy.[16-18]

At least nine diseases with expansions in tracts of CAG repeats encoding polyglutamine tracts in different genes have been found to lead to neurodegeneration and consequent neurological (and in some cases also psychiatric) symptoms.[19-21] The recent discovery that spinocerebellar ataxia 8 (SCA8) and Huntington's disease-like

2 (HDL2) might be caused by expression of a CAG/glutamine tract expansion on the antisense strand suggests that there may now be at least 11 polyglutamine diseases,[22,23] although the potential roles of RNA toxicity and other nonpolyglutamine pathogenic mechanisms in these diseases have not been ruled out.

Huntington's disease (HD) is the most common of these polyglutamine diseases, and was first described by George Huntington in 1872. HD is caused by a CAG repeat expansion encoding an extended polyglutamine tract in the huntingtin protein and ultimately leading to a triad of cognitive, psychiatric and motor symptoms.[24] HD has been the most intensely studied polyglutamine disease, leading to major new insights into how the CAG/glutamine repeat expansion leads to pathogenesis in specific brain areas as well as some peripheral organs.[25-27] The other polyglutamine diseases include six spinocerebellar ataxias (SCA1,2,3,6,7,17), as well as dentatorubral pallidolysian atrophy (DRPLA) and spinal and bulbar muscular atrophy (SBMA).[18,28] The potential molecular mechanisms whereby abnormally expanded polyglutamine tracts induce cellular toxicity are reviewed elsewhere.[20,21]

A fascinating recent discovery of non-ATG-initiated translation associated with at least some tandem repeat expansions,[29] has raised unexpected possibilities regarding mechanisms of pathogenesis. This finding, along with knowledge of antisense transcripts suggests that a single tract of DNA could generate at least two transcripts, seven reading frames and 'potentially nine toxic entities!'.[30]

A separate class of human disorders involve polyalanine expansions in specific proteins and mainly present as abnormalities of development.[31,32] These polyalanine disorders do not appear to be neurodegenerative, thus setting them apart from those diseases caused by expanded polyglutamine, but rather seem to be caused by a disruption of the normal function of the polyalanine tracts within the respective proteins.[32]

Other disorders of tandem repeat expansion include some nonpolyglutamine spinocerebellar ataxias, Friedreich ataxia, fragile X syndrome, fragile X tremor-ataxia syndrome (FXTAS) and myotonic dystrophy.[33-37] These disorders involve tandem repeats located in noncoding genomic regions and are therefore associated with abnormal gene expression, RNA structure and/or function.[18,38-40]

Collectively, these tandem repeat expansion disorders constitute a major personal, medical and economic burden, and therefore expanding our understanding and developing effective therapeutic approaches represents a clinical priority. Each disorder may require its own tailored therapeutic strategy, however it could be imagined that a therapy targeting polyglutamine toxicity, for example, may have efficacy across that wider group of neurodegenerative diseases. Furthermore, the study of these unique disorders, involving tandem repeat lengths beyond the normal range, may continue to provide insights into why tandem repeats have evolved, both in coding and noncoding regions of the human genome, and thus illuminate molecular and cellular processes in health and disease.

TANDEM REPEATS AS DYNAMIC MODULATORS OF DEVELOPMENT, BRAIN FUNCTION AND BEHAVIOR

Evidence has been provided for tandem repeats encoding homopeptide repeats in transcription factors as key regulators of developmental processes and subsequent anatomical diversity amongst various breeds of domestic dogs.[41] This study and other

evidence supports a role for homopeptide-repeat containing proteins in specific molecular and cellular processes, including transcriptional regulation.[1,7,42-44]

Tandem repeats, and their unique polymorphic contributions to genetic plasticity have been proposed to contribute to the modulation of brain development and function.[1] Putative roles for tandem repeats in various neurotransmitter and neuromodulatory systems, affecting brain development, behavioural modulation as well as affective and cognitive function have been recently reviewed.[2,16]

What is the evidence that tandem repeats are common in genes that are important for neural development and function? Firstly, the fact that the vast majority of known monogenic tandem repeat expansion disorders affect the nervous system provides indirect evidence, with the caveat being that many of those disease genes are also expressed in nonneural tissues. Nevertheless, the vulnerability of the nervous system to expansions of tandem repeats (both protein coding and noncoding) suggests that these tandem repeats may be particularly important in neural processes such as neurodevelopment and brain function. Secondly, bioinformatic approaches have indicated that many genes hosting tandem repeats are associated with neural functions.[1,45] Evidence for tandem repeats as modulators of behavioural processes has also been produced. For example, a TRP in the vasopressin receptor in prairie voles has been reported to modulate brain function and social behaviour.[46]

One implication of these and other studies is that TRPs could modulate normal human development, brain function, cognition and behavior.[1] A reasonable starting point for testing such a general hypothesis is to examine some of the many genes that have been implicated in monogenic human diseases.[16] The Huntington's disease (HD) gene, encoding a glutamine repeat in the huntingtin (Htt) protein as discussed above, provides one example. In the healthy range (nonHD families) Htt is known to have around 5-34 CAG/glutamine repeats. This TRP in Htt is known to be under selective pressure during evolution.[47] Furthermore, evidence from mice shows that Htt is important in development and brain function[48,49] and that removal of the CAG repeat has functional consequences.[50]

One specific hypothesis I therefore propose, as an extension of previous ideas,[1] is that healthy individuals who are polymorphic for their CAG/glutamine repeat in the Htt gene/protein will show differences in brain development, structure and function. A corollary of this would be that individuals who are 'gene-positive' for a CAG-expansion causing HD may have abnormal brain development, due to the functional effects of the long polyglutamine in Htt, prior to any 'toxic gain of function' leading to neurodegenerative changes. Consequently, in order to fully understand HD we will need to comprehend the role of the polyglutamine tract in the spatiotemporally regulated functions of Htt. This could be extended to many more tandem repeats, and their polymorphic variants, located in and around neurally expressed genes.

A PROPOSED ROLE FOR TANDEM REPEAT POLYMORPHISMS IN COMMON POLYGENIC DISORDERS

Recent genome-wide association (GWA) studies do not fully account for the major genetic contributions to common polygenic disorders, and this has led to an active search for the 'missing heritability'.[51,52] The evidence for tandem repeat expansions as

contributors to monogenic disorders of Mendelian inheritance has been outlined above. However, this may represent only a small fraction of tandem repeat contributions to human disease. It has been proposed that many common polygenic disorders may involve TRPs as major contributors to so-called 'missing heritability'.[2] TRPs could thus play a key role in modulating disease susceptibility for a range of common polygenic disorders, including various psychiatric and neurological disorders.[2] Recent evidence from studies of sporadic amyotrophic lateral sclerosis (ALS) supports the importance of TRPs as dynamic sources of genetic susceptibility in such complex polygenic diseases.[53-55]

The vast majority of GWA studies currently involve analysis of SNPs and therefore do not assay other major types of polymorphisms, such as TRPs. There are hundreds of thousands of distinct TRPs in the human genome, and this represents an extensive source of genetic variance possessing a dynamic and unique polymorphic range. SNPs are, with very few exceptions, binary in nature, whereas TRPs display extended digital distributions (multiallelic genotypes). An individual tandem repeat can thus exhibit a large array of polymorphic variants, which in turn extends the potential variety of genetic contributions to disease susceptibility.[2]

GWA studies will increasingly involve analyses of whole genome sequences, via the utilisation of next-generation sequencing technologies. It has been proposed that such new GWA studies involving whole-genome sequencing evaluate TRPs (using appropriate sequencing and bioinformatic protocols), as well as SNPs and other polymorphisms, for disease associations, so that the missing heritability may finally be found.[2] Rapid and accurate whole-genome sequencing has the potential to elucidate roles for TRPs in the dynamic modulation of biological evolution, development, function and dysfunction. It is therefore a priority to characterise the human 'repeatome', which I define here as the full set of tandem repeats and other repetitive DNA elements in the genome, as well as their transcribed and translated expression products. The repeatome would naturally encompass the full extent of TRPs throughout populations of a given species.

TANDEM REPEAT INSTABILITY AS A SOURCE OF CELLULAR HETEROGENEITY AND PHENOTYPIC DIVERSITY IN THE DEVELOPING AND MATURE ORGANISM

Instability of tandem repeats during meiosis, and consequent variability of repeat length between generations, is thought to be responsible for genetic anticipation in tandem repeat-expansion disorders such as HD. Meiotic instability of tandem repeat length could also contribute to common polygenic disorders, where variability in repeat lengths could modulate genetic susceptibility between generations.[2] Tandem repeats can also potentially change length during mitosis, and the most striking consequences of such repeat instability would be expected during development and may allow cellular selection to favour specific tandem repeat lengths in particular tissues and cell types.[1]

One additional potential source of tandem repeat length variability could involve postmitotic instability. Evidence for such tandem repeat instability in postmitotic neurons has recently been found.[56] Postmitotic instability of tandem repeats needs to be investigated in detail to establish its full extent, however it is already known that repeat instability (involving well described microsatellites) has been implicated in some sporadic cancers, which could result from postmitotic repeat length mutations.[57]

TANDEM REPEAT POLYMORPHISMS AND EVOLUTION

The abundance and broad distribution of tandem repeats in the human genome suggests that these sequences may arise randomly as 'genetic stutters' and that evolution actively selects for such repetitive DNA sequences.[3] Tandem repeats, and simple sequence repeats (SSRs) in particular, have been proposed to function as 'tuning knobs' during evolution.[3,12,16]

It has been observed that tandem repeat lengths are distributed in a digital manner, in contrast to the binary possibilities offered by SNPs; TRPs can thus act as 'digital genetic modulators' producing a continuously variable array of genotypes.[1] Furthermore, recent experimental evidence from a yeast model demonstrates that tandem repeats can enhance the 'evolvability' of specific promoter sequences.[58,59] Tandem repeats could thus be mediators of genetic plasticity, providing a diverse and dynamic genomic template, and associated phenotypic diversity, upon which natural selection may act.[1,12]

I propose a further possibility, that TRPs provide a dynamic source of stochastic genetic variation influencing organismal development, function and thus evolution. This would be analogous to the role of 'stochastic epigenetic variation' recently proposed by Feinberg and Irizarry.[60] More specifically, it was suggested that stochastic epigenetic variation might mediate phenotypic variability, without altering the mean phenotype, thus both enhancing evolutionary fitness whilst increasing susceptibility to disease of a population exposed to a changing environment.[60] These two hypotheses may in fact be interconnected, as evidence from repeat expansion disorders indicates that tandem repeats can modulate epigenetic processes, including chromatin remodelling.[34,39]

CONCLUSION

Tandem repeats are being increasingly recognised as a major potential source of genetic plasticity and associated cellular and organismal diversity, in humans and many other species.[1,3,61] Tandem repeat instability may mediate genomic plasticity and somatic variation, thus enhancing diversity at cellular, tissue and systems levels.[1] Tandem repeat length instability, occurring either meiotically, mitotically or postmitotically, thus has the capacity to mediate functional diversity at molecular, cellular, physiological and behavioural levels.

Technical and bioinformatic challenges need to be met to facilitate thorough analyses of tandem repeats and their polymorphic variants at a whole genome level.[62] Understanding genetic plasticity involving other types of repetitive DNA, such as specific retrotransposons,[63] will also be a priority. These approaches promise to shed new light on the genetic regulation of developmental, physiological and evolutionary processes. Furthermore, systematic characterisation of the 'human repeatome' may provide fundamental new insights into the genetic basis of disease.

ACKNOWLEDGMENTS

The NHMRC and an ARC Future Fellowship currently support the author's research. I thank past and present members of the Hannan Laboratory for useful discussions.

REFERENCES

1. Nithianantharajah J, Hannan AJ. Dynamic mutations as digital genetic modulators of brain development, function and dysfunction. Bioessays 2007; 29:525-535.
2. Hannan AJ. Tandem repeat polymorphisms: modulators of disease susceptibility and candidates for 'missing heritability'. Trends Genet 2010; 26:59-65.
3. King DG. Evolution of simple sequence repeats as mutable sites. In: Hannan AJ, ed. Tandem Repeat Polymorphisms: Genetic Plasticity, Neural Diversity and Disease. Austin/New York: Landes Bioscience/ Springer Science+Business Media, 2012:10-25.
4. Sawaya SM, Bagshaw ATB, Buschiazzo E et al. Promoter microsatellites as modulators of human gene expression. In: Hannan AJ, ed. Tandem Repeat Polymorphisms: Genetic Plasticity, Neural Diversity and Disease. Austin/New York: Landes Bioscience/Springer Science+Business Media, 2012:41-54.
5. Hamada H, Seidman M, Howard BH et al. Enhanced gene expression by the poly(dT-dG) poly(dC-dA) sequence. Mol Cell Biol 1984; 4:2622-2630.
6. Kashi Y, King D, Soller M. Simple sequence repeats as a source of quantitative genetic variation. Trends Genet 1997; 13:74-78.
7. Faux NG, Bottomley SP, Lesk AM et al. Functional insights from the distribution and role of homopeptide repeat-containing proteins. Genome Res 2005; 15:537-551.
8. Verstrepen KJ, Jansen A, Lewitter F et al. Intragenic tandem repeats generate functional variability. Nat Genet 2005; 37:986-990.
9. Roberts J, Scott AC, Howard MR et al. Differential regulation of the serotonin transporter gene by lithium is mediated by transcription factors, CCCTC binding protein and Y-Box binding protein 1, through the polymorphic intron 2 variable number tandem repeat. J Neurosci 2007; 27:2793-2801.
10. Salichs E, Ledda A, Mularoni L et al. Genome-wide analysis of histidine repeats reveals their role in the localization of human proteins to the nuclear speckles compartment. PloS Genet 2009; 5:e1000397.
11. Mattick JS, Taft RJ, Faulkner GJ. A global view of genomic information—moving beyond the gene and the master regulator. Trends Genet 2010; 26:21-28.
12. Kashi Y, King DG. Simple sequence repeats as advantageous mutators in evolution. Trends Genet 2006; 22:253-259.
13. Richard GF, Kerrest A, Dujon B. Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. Microbiol Mol Biol Rev 2008; 72:686-727.
14. Conrad DF, Pinto D, Redon R et al. Origins and functional impact of copy number variation in the human genome. Nature 2010; 464:704-712.
15. Richards RI. Dynamic mutations: a decade of unstable expanded repeats in human genetic disease. Hum Mol Genet 2001; 10:2187-2194.
16. Fondon JW III, Hammock EA, Hannan AJ et al. Simple sequence repeats: genetic modulators of brain function and behavior. Trends Neurosci 2008; 31:328-334.
17. Usdin K. The biological effects of simple tandem repeats: lessons from the repeat expansion diseases. Genome Res 2008; 18:1011-1019.
18. Van Eyk CL, Richards RI. Dynamic mutations: where are they now? In: Hannan AJ, ed. Tandem Repeat Polymorphisms: Genetic Plasticity, Neural Diversity and Disease. Austin/New York: Landes Bioscience/ Springer Science+Business Media, 2012:55-77.
19. Bauer PO, Nukina N. The pathogenic mechanisms of polyglutamine diseases and current therapeutic strategies. J Neurochem 2009; 110:1737-1765.
20. Polling S, Hill AF, Hatters DM. Polyglutamine aggregation in Huntington and related diseases. In: Hannan AJ, ed. Tandem Repeat Polymorphisms: Genetic Plasticity, Neural Diversity and Disease. Austin/New York: Landes Bioscience/Springer Science+Business Media, 2012:125-140.
21. Robertson AL, Bottomley SP. Molecular pathways to polyglutamine aggregation. In: Hannan AJ, ed. Tandem Repeat Polymorphisms: Genetic Plasticity, Neural Diversity and Disease. Austin/New York: Landes Bioscience/Springer Science+Business Media, 2012:115-124.
22. Moseley ML, Zu T, Ikeda Y et al. Bidirectional expression of CUG and CAG expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8. Nat Genet 2006; 38:758-769.
23. Wilburn B, Rudnicki DD, Zhao J et al. An antisense CAG repeat transcript at JPH3 locus mediates expanded polyglutamine protein toxicity in Huntington's disease-like 2 mice. Neuron 2011; 70:427-440.
24. Waldvogel H, Thu D, Hogg V et al. Selective neurodegeneration, neuropathology and symptom profiles in Huntington's disease. In: Hannan AJ, ed. Tandem Repeat Polymorphisms: Genetic Plasticity, Neural Diversity and Disease. Austin/New York: Landes Bioscience/Springer Science+Business Media, 2012:141-152.
25. Zuccato C, Valenza M, Cattaneo E. Molecular mechanisms and potential therapeutical targets in Huntington's disease. Physiol Rev 2010; 90:905-981.