

TAR Cloning: Perspectives for Functional Genomics, Biomedicine, and Biotechnology

Natalay Kouprina¹ and Vladimir Larionov¹

¹Developmental Therapeutics Branch, National Cancer Institute, Bethesda, MD 20892, USA

Completion of the human genome sequence and recent advances in engineering technologies have enabled an unprecedented level of understanding of DNA variations and their contribution to human diseases and cellular functions. However, in some cases, long-read sequencing technologies do not allow determination of the genomic region carrying a specific mutation (e.g., a mutation located in large segmental duplications). Transformation-associated recombination (TAR) cloning allows selective, most accurate, efficient, and rapid isolation of a given genomic fragment or a full-length gene from simple and complex genomes. Moreover, this method is the only way to simultaneously isolate the same genomic region from multiple individuals. As such, TAR technology is currently in a leading position to create a library of the individual genes that comprise the human genome and physically characterize the sites of chromosomal alterations (copy number variations [CNVs], inversions, translocations) in the human population, associated with the predisposition to different diseases, including cancer. It is our belief that such a library and analysis of the human genome will be of great importance to the growing field of gene therapy, new drug design methods, and genomic research. In this review, we detail the motivation for TAR cloning for human genome studies, biotechnology, and biomedicine, discuss the recent progress of some TAR-based projects, and describe how TAR technology in combination with HAC (human artificial chromosome)-based and CRISPR-based technologies may contribute in the future.

The ability to quickly and accurately isolate full-length genes, including coding (exons) as well as non-coding regions (introns and up- and downregulators) that participate in controlling gene expression, has become especially relevant in recent years in light of the growing interest in gene function studies and gene therapy. Gene therapy is a technique for correcting defective genes responsible for disease development.^{1–4} Researchers identify a defect in a gene that is responsible for a disease and attempt to replace the defective gene with a healthy one, either by using viral and non-viral approaches that are based on expression of full-length cDNA in gene-deficient cells⁵ or by repairing the defect using CRISPR/Cas9 genome editing systems.⁶ The field of genetics is also becoming increasingly important in drug design for personalized medicine. Being able to monitor the effect of a drug treatment on a genetic level greatly enhances our understanding of how a drug affects the patient. When the effect of the drug is better understood, this knowledge can be

used to create more effective versions of that treatment and to create variants of the drug that are more effective for specific groups of patients.^{7,8} The difficulty of isolating and manipulating gene variants has been a great hindrance to these areas of research.

The transformation-associated recombination (TAR) cloning approach allows entire genes and large chromosomal regions to be selectively and accurately isolated from total genomic DNA.^{9–27} The method exploits a high level of recombination between homologous DNA sequences during transformation in the yeast *S. cerevisiae*. **Figure 1** illustrates TAR isolation of a gene or locus from a genome. Recombination between the homologous sequences from a genome and targeting sequences from the TAR vector leads to establishment of a circular molecule in a YAC (yeast artificial chromosome) form containing the desired genomic fragment carrying a gene or locus. When TAR cloning is applied to complex genomes, such as the human genome, the technique produces YAC clones containing the desired genomic fragment at a frequency of 0.5%–2% of all yeast colonies screened. With the optimized protocol, which includes pre-treatment of genomic DNA with CRISPR/Cas9 nucleases to generate double-strand breaks near the targeted genomic region, the yield of gene- or region-positive clones may be as high as 32%.²⁸ Because the same TAR vector can be simultaneously applied to DNA samples isolated from multiple individuals, the TAR cloning approach may be used to generate a bank of gene variants from different population groups (**Figure 1**) for gene functional studies and for mutational analysis of the same gene using DNA samples from patients with different diseases.

At present, TAR cloning has become a valuable procedure for manipulation of large DNA molecules. The ability to selectively isolate genomic regions and genes has greatly advanced structural and functional analysis of complex and simple genomes and also has different practical applications for biomedicine and biotechnology.²⁷ Over the past years, TAR cloning has been used to isolate dozens of unique regions and full-length genes from the genomes of humans, African

<https://doi.org/10.1016/j.omtm.2019.05.006>

Correspondence: Natalay Kouprina, Developmental Therapeutics Branch, National Cancer Institute, Bethesda, MD 20892, USA.

E-mail: kouprinn@mail.nih.gov

Correspondence: Vladimir Larionov, Developmental Therapeutics Branch, National Cancer Institute, Bethesda, MD 20892, USA.

E-mail: larionov@mail.nih.gov



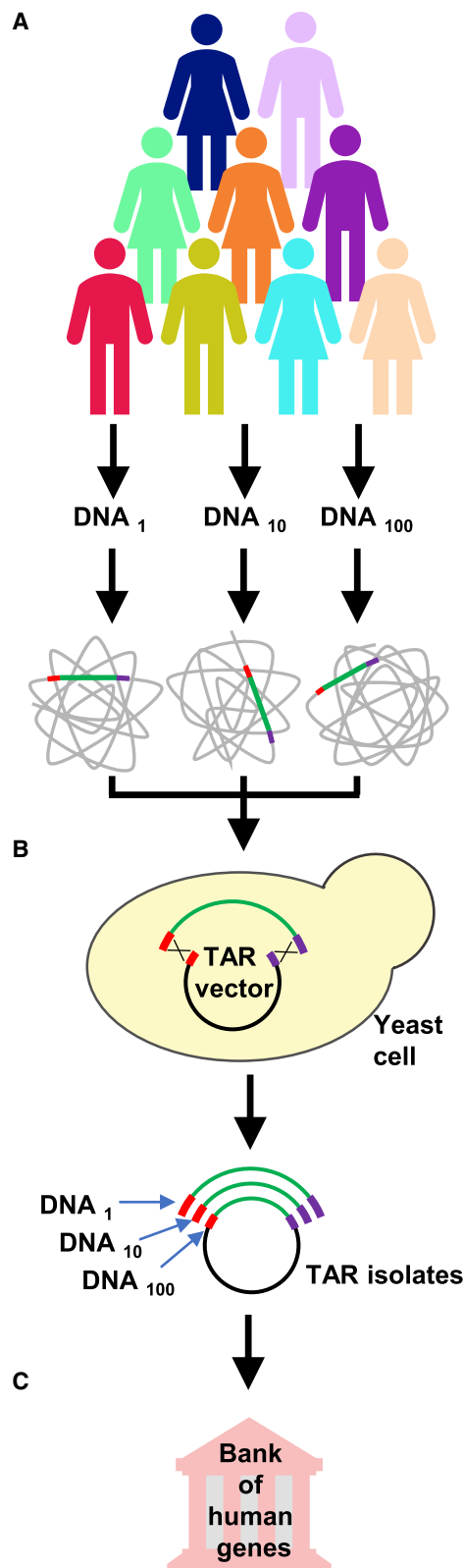


Figure 1. Isolation of a Specific Region or Gene from Multiple Human Genomes by TAR Cloning in Yeast *S. cerevisiae*

(A) The TAR cloning procedure includes preparation of yeast spheroplasts and transformation of the spheroplasts by gently isolated high-molecular-weight total genomic DNA along with a TAR vector containing the targeting sequences (hooks). The hooks can be as small as 60 bp.¹¹¹ (B) For TAR cloning experiments, the vector DNA is linearized by a unique endonuclease located between the hooks to expose targeting sequences. Recombination between the vector hooks and homologous sequences in the co-transformed human DNA fragment results in rescue of the desired region as a circular yeast artificial chromosome (YAC) that replicates and segregates properly.^{11–42} The same TAR vector can be applied to multiple DNA samples. If genomic DNA is pre-treated with CRISPR/Cas9 nucleases to generate double-strand breaks near the targeted genomic region, this results in a dramatic increase in the fraction of gene-positive colonies (up to 32 times).²⁸ Because the TAR vector contains a BAC cassette, a TAR-isolated YAC clone containing a gene or genomic fragment of interest can be easily moved to bacterial cells by electroporation. (C) Using a set of TAR vectors with different hooks, TAR cloning may help construct a bank of full-length human genes ready to be transferred into human cells for gene function studies and, potentially, for gene therapy.

great apes, the mouse, and other organisms (Table 1).^{11–15,18–23,29–32} Application of the TAR cloning strategy has been demonstrated for long-range haplotyping,^{33–37} and, as a result, this technique is well suited for large-scale analysis of haplotypes in multiple heterozygous individuals to identify haplotypes that might contribute to disease. Because up to 15% divergence in DNA sequences does not prevent cloning in yeast,³⁸ TAR technology is a powerful tool for comparative genomics and evolutionary studies. Many gene homologs from nonhuman primates have been TAR-isolated using available human sequence information for targeting.^{29–32} TAR cloning may be put to use in the final phase of complex genome sequencing and significantly contribute to closing the gaps and verifying the accuracy of contig assembly.^{39–42} Many genes in the human genome reside within segmental duplications (SDs) with a high level of similarity, which prevents their mutational analysis by routine PCR methods. A TAR cloning strategy has been successfully applied to recover and analyze several duplicated regions, and, therefore, this approach is applicable to all gene families whose analysis is prohibited by the presence of SDs.^{29,32,34–36} Accumulating information shows a plethora of structural variations (such as deletions, insertions, inversions, translocations, and copy number variations [CNVs]) in genomes of different individuals, some of which may lead to disease.^{43–47} TAR cloning thus becomes a unique tool to discover and characterize the structural variations causing Mendelian disorders.^{35–37} This information may be valuable for diagnostics as well as for the development of new drugs. As seen from the past decade, the cost, quality, and efficacy of the TAR cloning method has been improved significantly by rigorous and comprehensive testing for the accuracy of locus isolation from different organisms, shifting the technology from being a luxury to becoming essential and routine.

TAR Cloning to Provide a Collection of Full-Length Human Genes

At present, the science of genomes has moved from a phase of genome sequencing into a phase of analysis of gene expression and function. Conventional gene therapy has some success in curative treatment for some primary immune deficiencies in

**Table 1. Examples of Chromosomal Regions and Genes or Gene Clusters Isolated by TAR Cloning**

Gene or Region	Source	Location	Size (kb)
<i>BRCA2</i>	human	13q13.1	100
<i>BORIS</i>	human	20q13.31	80
<i>TERT</i>	human	5p15.33	150
<i>SPANX-D</i>	human	Xq27.2	107
<i>PTEN</i>	human	10q23.31	110
<i>TM</i>	human	3p25.3	50
<i>NBS1</i>	human	8q21.3	60
<i>PKD1</i>	human	16p13.3	74
Locus <i>CAD</i>	human	9p21.3	230
<i>KAI1</i>	human	11p11.2	180
Gap1	human	19p13.1	90 ^a
Gap 2	human	19p13.2	120 ^a
Gap3	human	19p13.3	180 ^a
Gap6	human	19p13.4	100 ^a
<i>HPRT</i>	human	Xq26.1	250
	human	17q21.31	84 ^b
	chimpanzee	17	84
	gorilla	nd	66
	orangutan	nd	65
	rhesus macaque	nd	65
	human	1q31	65 ^b
	chimpanzee	1	65
<i>ASPM</i>	gorilla	nd	65
	orangutan	nd	65
	rhesus macaque	nd	65
	human	Xq27.2	83 ^b
	chimpanzee	X	58
	gorilla	nd	61
	bonobo	nd	58
	human	Xq27.1	90 ^b
<i>SPANX-B</i>	chimpanzee	X	63
	gorilla	nd	63
	human	8q21.3	90 ^b
<i>NBS1</i>	chimpanzee	8	60
	bonobo	nd	60
	human	11q22	150 ^b
<i>ATM</i>	rhesus macaque	nd	140
<i>Hpii</i>	mouse	X	200
<i>Peg3</i>	mouse	7	180
<i>Muc2</i>	mouse	7	100
Locus <i>RNR4</i>	human	21.pl2	820 ^c
Xq27	human	Xq27.1-3	750 ^d
Gene cluster	<i>Saccharomonospora</i> sp.		67

(Continued)

Table 1. Continued

Gene or Region	Source	Location	Size (kb)
Gene cluster	<i>Salinispora pacifica</i>		21
Gene cluster	<i>Bacillus subtilis</i>		17
Gene cluster	<i>Pseudoalteromonas piscicida</i>		34
Gene cluster	<i>B. subtilis</i>		38
Gene cluster	<i>S. albus</i>		36
Gene cluster	<i>S. aureofaciens</i>		44
Gene cluster	<i>S. sclerotialis</i>		33

nd, not determined.

^aSequences missing on the chromosome 19 were TAR-cloned, sequenced, and deposited into GenBank.³⁹^bThe same regions or genes were cloned from different species using TAR vectors containing human-specific targeting sequences developed from 5' and 3' gene-flanking regions.²⁹⁻³²^cMultiple clones carrying human rDNA units with sizes up to 140 kb were isolated from chromosome 21. Accurate long-read sequencing of 13 TAR isolates of rDNA units covers ~0.82 Mb of the chromosome 21 rDNA cluster.⁴²^dDirect isolation of a set of overlapping 50- to 120-kb genomic segments by *in vivo* recombination in yeast was used to perform a mutational analysis of the 750-kb region in X-linked families. These results excluded the 750-kb genetically unstable region at Xq27 as a candidate locus for prostate malignancy.³⁶

the last decade because of advances in genome editing.⁴⁸ More specifically, development of the CRISPR system created the opportunity to site-specifically correct mutated DNA base pairs or insert a corrective cDNA minigene while maintaining gene expression under the control of endogenous regulatory elements.^{6,49,50}

Nevertheless, although cDNA clones are still widely used for gene expression studies, the availability of clones containing full-length genes with all of the necessary regulatory regions, including their own promoter and intronic sequences, would provide indisputable advantages because only in such a configuration can “physiological” expression be achieved. Until recently, the only sources of full-length genes were clones from BAC (bacterial artificial chromosome) and YAC libraries. However, construction of gene expression cassettes using BAC or YAC clones, if found, is time consuming and hit or miss because a gene is often available as a set of fragments in different clones that must be laboriously pieced together. Even if a yeast or BAC contains an entire gene, its use for expression experiments requires trimming of the insert to remove unrelated flanking sequences, often including extraneous genes. In contrast, TAR cloning technology is selective and efficient and allows high-fidelity isolation of full-length genes with all of their regulatory elements (in these cases, the TAR-isolated human genes and their homologs were sequenced).^{19,22,27,29-31} In perspective, TAR cloning engineering may provide a complete collection of full-length human genes (Figure 1) ready for CRISPR-mediated insertion or replacement with the following gene expression studies to elucidate the biological function of the disease-associated genes and to help with gene therapy for disease treatment.

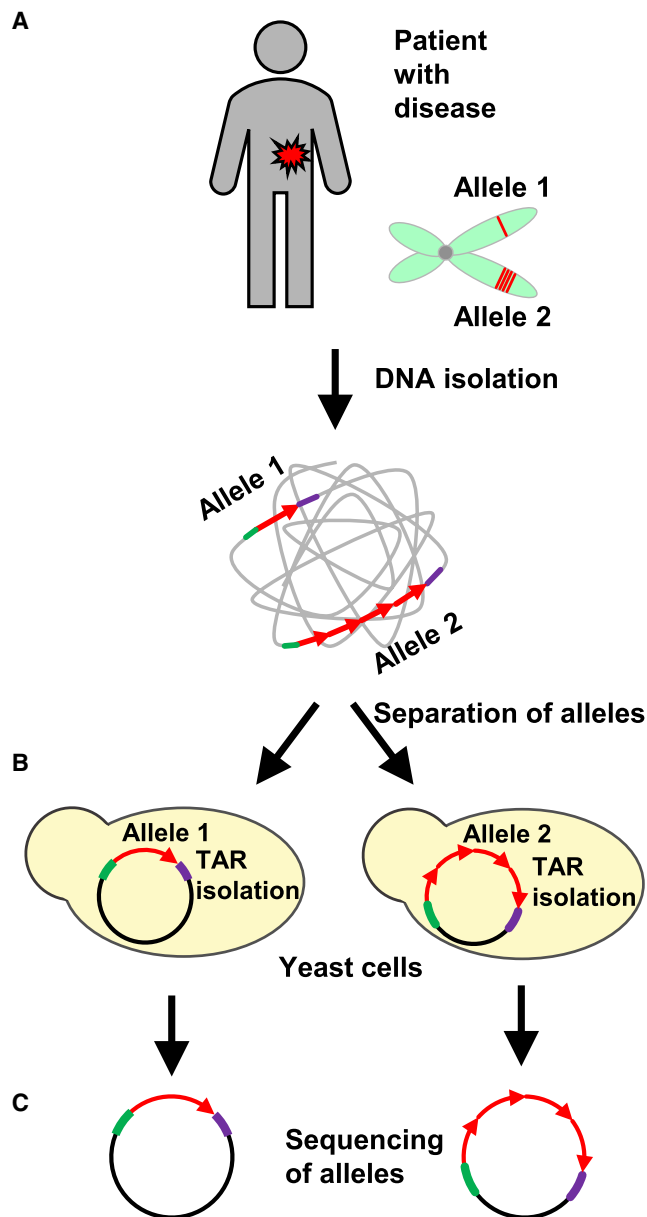


Figure 2. Separation of Parental Alleles by TAR Cloning

(A) TAR cloning may be employed to isolate parental alleles of any gene from a patient^{33–37} or a normal individual and determine the specific mutations or rearrangements at polymorphic sites of each allele. The diagram shows separation and isolation of two alleles, allele 1 (normal) and allele 2 (mutant), from the unique genomic region of a patient. (B) Yeast spheroplasts are transformed with genomic DNA isolated from the patient, along with a linearized TAR cloning vector containing 5' and 3' unique sequences (green and in purple) specific to the targeted region (red). (C) After TAR cloning in yeast and transfer of TAR isolates to bacterial cells, BAC DNAs are sequenced to identify rearrangements/mutations specific to an allele.

TAR Cloning to Study the Genetic Basis of Human Disease

It is well known that hereditary genomic alterations frequently underlie genetic predisposition to disease in humans, including the initia-

tion and progression of cancer. Genomic loss or amplification, translocations, and inversions in many specific regions show consistent association with certain diseases, indicating that certain regions harbor key disease-associated alterations.^{28,29,31,32,38} Importantly, the vast majority of such alterations remain uncharacterized, in part because of technical hurdles. Systematic identification of these chromosomal rearrangements and their cloning and characterization at the molecular level are increasingly needed to obtain a comprehensive description of the genomic alterations leading to cancer or age-related disease. Such an effort would provide the most important and general foundation for basic, clinical, and commercial work in the future. Comprehensive identification of genomic alterations predisposing to disorders requires examination of a substantial number of human individuals. In principle, third-generation sequencing technologies, such as the Oxford Nanopore producing reads up to 700 kb, could be applied for this study.^{51,52} However, wide use of long-read sequencing is limited by its high error rate (approximately 2 of 10 bases).^{51,52} Thus, achieving this ambitious goal is within reach by using TAR cloning technology, which provides a unique tool to rapidly isolate the same region or gene with a certain allele^{33–37} (Figure 2) from hundreds of individuals or patients to identify and characterize a specific chromosomal variation linked to disease. A good example of the application of TAR cloning for verification of gene disease mapping is the analysis of the Xq27-q28 region. Previously, several linkage studies have provided evidence for the presence of the hereditary prostate cancer locus, HPCX1, at Xq27-q28.⁵³ The strongest linkage peak for prostate cancer overlaid a variable region of ~750 kb at Xq27 enriched by large SDs, suggesting that predisposition to prostate cancer may be a genomic disorder caused by recombinational interaction between SDs. Direct isolation of a set of overlapping genomic segments carrying SDs completely covering the 750-kb region by TAR cloning was used to perform a mutational analysis of this region in patients. The subsequent analysis of these TAR-isolated fragments excluded the 750-kb genetically unstable region at Xq27 as a candidate locus for prostate malignancy.³⁶ The polymorphism information accumulated during analysis of this region may be useful for the analysis of other disease-associated loci that were also mapped to Xq27-q28, including the TGCT locus linked to hereditary testicular cancer and the susceptibility loci for dyslexia, autism, and migraine.^{54,55} In the future, the accumulated information from individual genomes can help predict the pharmaceutical drug response for each individual and expand the therapeutic window of a treatment. In addition, the uniqueness of the technology will help with the design of diagnostics for genomic disorders caused by chromosomal rearrangements.

TAR Cloning to Illuminate Genomic Dark Matter

The recently reported human genome sequence (hg38) still includes some gaps. For example, five acrocentric chromosomes contain gaps corresponding to clusters of rRNA (rDNA) genes. Despite the key role of the human ribosome in protein biosynthesis, little is known about the extent of rDNA sequence variation within one chromosome and between chromosomes. Prior attempts to characterize rDNA tandem repeat regions have failed because of their highly

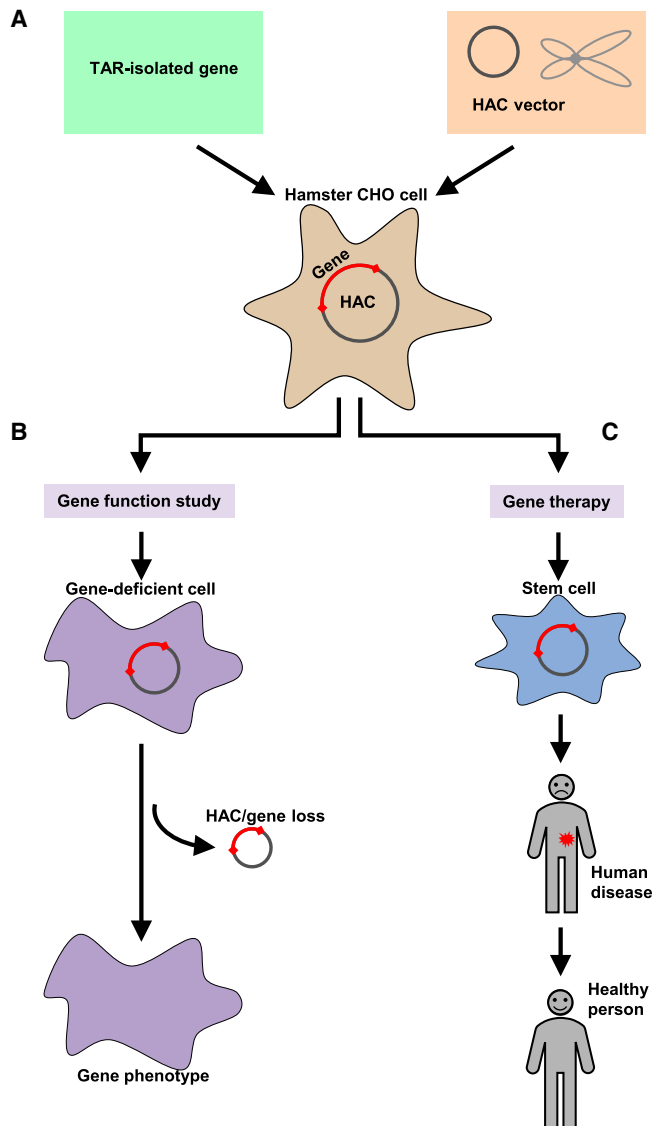


Figure 3. Schematic Diagram Illustrating the Possible Spectrum of Applications of the HAC-Based Gene Delivery Vector Carrying a TAR-Isolated Human Gene for Pluripotent Stem Cell-Based Regenerative Medicine and Gene Function Studies

(A) A full-length TAR-isolated gene (green box) is loaded into the HAC-based vector via the Cre/loxP-mediated recombination in hypoxanthine phosphoribosyltransferase (HPRT)-deficient Chinese hamster ovary (CHO) cells. (B) The HAC, along with a gene of interest, is then delivered to target cells (for example, mouse embryonic stem cells [ESCs] or patient-derived human cells deficient for this gene) via an improved microcell-mediated chromosome transfer (MMCT) procedure.^{112–114} Thus, the HAC containing a gene can be utilized for functional analyses *in vitro* and *in vivo*, including a humanized mouse model. For the HAC with a regulated kinetochore, the phenotypes arising from stable gene expression from the HAC can be reversed when the HAC is eliminated from the cells by inactivating its kinetochore, which provides control over the phenotypic changes attributed to expression of HAC-encoded genes. (C) For gene therapy, a HAC containing a therapeutic gene can be utilized for treatment of patients with genetic disorders. As a first step, induced pluripotent stem cells (iPSCs) should be pro-

repetitive structure. Recently, human chromosome 21 rDNA segments were recovered using TAR cloning technology.⁴² Analysis of these TAR-isolated clones revealed previously missed palindromic structures as well as hundreds of variant alleles in the 45S transcription unit and in the non-transcribed spacer. The large number of variants reveals more—and more universal—heterogeneity in human rDNA than considered previously, opening the possibility of corresponding variations in ribosome dynamics. The variants can be assessed for any genetic associations, and the TAR clones provide reagents for further study of human nucleolar assembly and activity. Different variants of human rDNA units may help to clarify the peculiarity of PolII transcription machinery as well as the mechanism by which rDNA units are selected and targeted for chromatin changes leading to heterochromatinization in nucleolar organizer regions (NORs). It is interesting that some variants of a coding region of 21 rDNA units recovered by TAR cloning are located in segments that are conserved between human and yeast.²⁶ Functional analysis of these variants *in vivo* (i.e., in the presence of many copies of rRNA genes in the chromosomal rDNA repeats) is possible using the yeast rDNA assay developed by Chernoff et al.⁵⁶ Thus, sequence variants identified in 18S and 28S regions that are conserved between human and yeast can be analyzed. In the future, TAR cloning may be applied to cover the rDNA gaps on other chromosomes.

Knowledge of rDNA function is very important because of the dosage effects of human ribosomal genes (rDNA) in health and disease.⁵⁷ rRNA gene dosage affects normal growth and aging, stress resistance of healthy individuals, and survivability of patients with chromosomal abnormalities as well as the risk and severity of some multifactorial diseases with proven genetic predisposition. A study showed that patients with acute systemic purulent infection who had developed multiple organ failure had a low genomic dosage of active ribosomal genes.⁵⁸ The authors concluded that the pattern of purulent inflammation is more severe in subjects with low genomic dosages of active ribosomal genes. Reduced ribosomal activity has been suggested to be a major contributor to the pathology of Alzheimer's disease (AD).⁵⁹ The rDNA copy numbers were also higher in SZ (schizophrenia) patients.⁶⁰ The presence of palindromic structures as well as certain variant alleles in the 45S transcription may also contribute to human diseases. Thus, TAR cloning of gaps corresponding to clusters of rRNA genes and their subsequent analysis is very important for elucidating the source of and predisposition to many diseases.

TAR-Isolated Genes Expressed from HAC-based Delivery Vectors for Gene Function Studies with a Potential for Gene Therapy

Despite progress with the CRISPR system, which is regarded as an eminent genome engineering tool in therapeutics,^{6,50} we still believe that further progress in functional genomics and gene therapy

duced from patient fibroblasts¹¹⁴ using one of the currently available protocols. Then a HAC carrying a therapeutic gene is introduced into iPSCs^{115,116} from donor CHO cells using MMCT. The final step would include transplantation of stem cells into the patient.

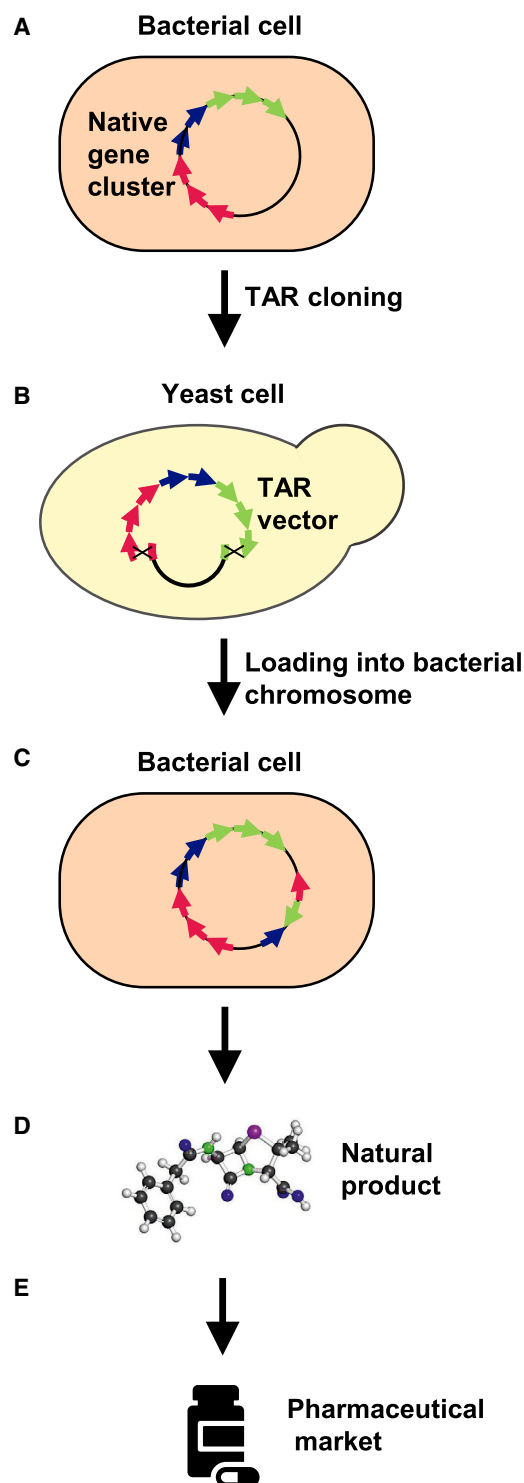


Figure 4. Design of TAR Cloning of Large Natural Product Biosynthetic Gene Clusters from Simple Genomes

(A) Physical map of a gene cluster located on a chromosome propagated in a bacterial cell. In a TAR vector, the hooks or homology arms correspond to both ends of the cluster. (B) TAR-based natural product discovery involves homologous

depends on the availability of both full-length genes and a suitable system for gene delivery and gene expression. The most promising vector system is now human artificial chromosomes (HACs).^{61–69} HACs represent extra chromosomes carrying all of the required components of a functional kinetochore. HACs have advantages as gene expression vectors with potential use in gene therapy. They are stably maintained at a low copy in the host nucleus. Furthermore, they contain no viral genes or proteins and, therefore, should not cause severe immunogenic responses, which have been found to be a serious problem with adenoviral vectors.^{61–69} HACs are thus intrinsically safer than traditional viral gene therapy systems. They also have a large genetic packaging capacity, up to several megabases, that is suitable for carrying intact mammalian genes in the context of all of their long-range controlling elements that should confer physiological levels of fully regulated gene expression. Existing strategies for the creation of HACs can be broadly divided into two classes: top down, based on truncation of an existing chromosome into a much smaller mini-chromosome suitable for further manipulation^{63,64}, and bottom up, where defined, cloned chromosomal elements (alphoid or alpha-satellite units) are TAR-assembled into a prefabricated molecule^{70,71} that is capable of nucleating formation of a HAC *de novo* upon its introduction into human cells.^{70,72} Several laboratories have demonstrated the efficacy of HACs for delivery and expression of full-length genes up to megabase size *in vitro* (human cell lines) and *in vivo* (mice).^{73–82}

At present, HAC vectors are optimized for targeting the genomic regions isolated by TAR technology: (1) gene expression from the HAC-based vector is protected by chromatin insulators,⁸³ and (2) an unlimited number of genomic DNA fragments or genes may be assembled on the HAC molecule.^{84–86} Combining engineered human chromosomes with TAR cloning technology can greatly advance gene function studies (Figure 3).^{61,62,65,73,74} Moreover, the combination of large-capacity vectors carrying TAR-isolated genes with stem or progenitor cells establishes a novel platform for gene therapy.

TAR Cloning of Natural Product Biosynthetic Gene Clusters for Biotechnology

Over the past years, there have been many examples of direct capture of genes or gene clusters by TAR cloning from simple genomes; for example, recovery of large natural product (NP) biosynthetic gene clusters (Figures 4A and 4B).^{87–101} So far, TAR cloning has been applied for isolation of genomic regions and genes up to ~250 kb (Table 1). Therefore, the TAR method has an advantage, especially in the case of NP biosynthetic gene clusters in microbes, whose size is much smaller. The TAR-cloned natural genes or gene clusters

recombination between the linearized pathway capture TAR vector and co-transformed genomic DNA isolated from host bacterial cells that leads to formation of a circular YAC/BAC construct. (C and D) Some TAR-cloned natural genes or gene clusters may be transferred to a bacterial host strain and, if necessary, integrated into the chromosome of a host strain for basic research or production of natural compounds. (E) Natural products have provided considerable value to the pharmaceutical industry over the past half-century.

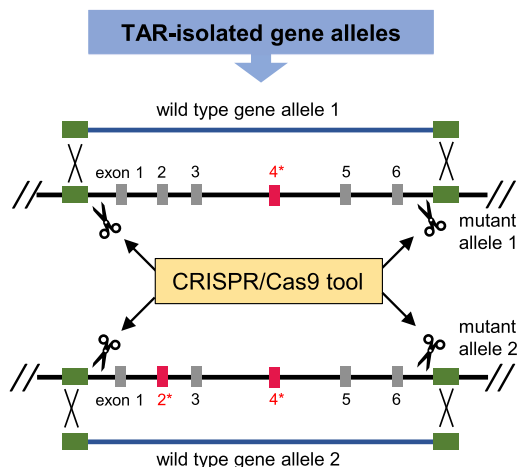


Figure 5. Schematic of a CRISPR/Cas9-Based Strategy to Simultaneously and Differentially Replace Two Mutant Gene Alleles of a Diploid Cell by Two TAR-Cloned Wild-Type Gene Alleles in a Single Step

A gene of interest is targeted with two single guide RNA (sgRNAs; green rectangles) at the 5' and 3' ends of the gene. Two repair templates are introduced into the cell, which mediate the replacement of the gene on both alleles. Mutant exons are marked in red.

may be transferred from yeast to a bacterial host strain as BACs-YACs or, if necessary, integrated into the chromosome of a host strain for basic research or production of natural compounds (Figures 4C and 4D). More specifically, the lipopeptide gene cluster from the marine actinomycete *Saccharomonospora* sp., which yields the antibiotic taromycin A,⁸⁹ the *Salinispora pacifica* NP gene cluster for biosynthesis of enterocin,⁹⁰ and the amicoumacin biosynthetic gene cluster from the marine *Bacillus subtilis* 1779 for antibiotic production in the *Bacillus* host⁹¹ were selectively TAR isolated. NPs and their derivatives contribute to a huge pharmaceutical market share comprising 61% of anticancer drugs and 49% of anti-infection medicine in the past 30 years¹⁰².

It is worth noting that TAR cloning requires that the cloned DNA fragment carries at least one autonomously replicating sequence (ARS) that can function as the origin of replication in yeast. Sequences that function as ARSs in yeast occur, on average, every 20–40 kb in all of the eukaryotic genomes examined so far,^{13,17} which suggests that most mammalian chromosomal regions can be readily isolated by TAR cloning with a vector that lacks an ARS. The ARS frequency in some microbe genomes may be reduced, meaning some regions may not be isolatable by the original TAR method. Therefore, we developed a general TAR cloning system for isolation of genomic regions regardless of the presence of ARS elements.¹⁰³ In this system, the TAR vector contains an ARS and a counter-selectable marker. Negative genetic selection eliminates the background because of vector recircularization caused by end joining during yeast transformation. Recently, a similar, very effective TAR capture from microbe genomes was developed by another lab⁹⁸ that also used a counter-selectable marker to improve efficiency of cloning. Both methods

are powerful tools for structural and functional analysis of prokaryotic genomes.

Isolation and manipulation of NP biosynthetic gene clusters have further accelerated our understanding of their molecular biosynthetic mechanisms, leading to a rational redesign of novel NPs.¹⁰⁴ To summarize, TAR cloning of gene clusters from microbe genomes provides an opportunity to address fundamental problems in environmental evolution and biotechnology.

Conclusions

TAR engineering may physically characterize all of the sites of chromosomal alterations in the human genome associated with predisposition to different diseases, including cancer. Accumulated comprehensive knowledge of the genetic basis of these diseases will provide a foundation for future medical research and have far-reaching implications for basic, clinical, and commercial efforts to understand, prevent, and treat diseases and develop new strategies for their diagnosis and treatment. TAR technology may become a foundation for creating a collection of annotated human genes, each of which is represented by a genomic copy containing coding regions (exons), 5' upstream and 3' downstream regulatory sequences, and introns. The availability of full-length genes with their own intact regulatory elements will catalyze major breakthroughs in functional, structural, and comparative genomics; diagnostics; gene replacement; and generation of animal models for human diseases, with a potential for gene therapy from which a new generation of genomic research products could be launched. Notably, the large size of genes or gene clusters does not prevent their propagation in yeast based on TAR assembly of whole-microbe genomes with sizes greater than 1.5 Mb.^{105–107} This unique technology may be used in combination with HAC-based gene delivery vectors for stable and regulated production of therapeutic proteins through regulated expression of full-length human genes. In addition, these technologies should improve the success rate of drug development related to efficacy and safety for humans.¹⁰⁸ In perspective, given the potential of engineered human chromosomes of curing genetic defects, HACs offer tremendous potential for gene therapy applications. Development of a universal gene delivery system for conventional gene therapy and for stable and regulated production of therapeutic proteins may be an ambitious goal in the near future.

For the past decade, CRISPR/Cas9-based methodologies have promise to lead to rapid manipulation of endogenous genes that will open the door for new synthetic biology platforms. Recently, a new gene replacement strategy utilizing CRISPR/Cas9 technology was devised.^{109,110} In this strategy, a mutant gene in the human genome is simultaneously and differentially replaced on both alleles. CRISPR/Cas9 technology combined with TAR-cloned gene alleles may be applied for gene functional analysis with perspectives on gene therapy (Figure 5). When such an approach is optimized, the study of complex biological systems will become more powerful and convincing for the investigator. In future, TAR-based gene technology may be used to carry out a regular customer service for isolation of genomic



regions and entire loci from clinical materials and alternative genomes (Table 1).

CONFLICTS OF INTEREST

The authors declare no competing interests.

ACKNOWLEDGMENTS

This work was supported by the Intramural Research Program of the NIH, the National Cancer Institute, and the Center for Cancer Research.

REFERENCES

- Salzman, R., Cook, F., Hunt, T., Malech, H.L., Reilly, P., Foss-Campbell, B., and Barrett, D. (2018). Addressing the Value of Gene Therapy and Enhancing Patient Access to Transformative Treatments. *Mol. Ther.* 26, 2717–2726.
- Jeevanandam, J., Pal, K., and Danquah, M.K. (2019). Virus-like nanoparticles as a novel delivery tool in gene therapy. *Biochimie* 157, 38–47.
- Mingozzi, F., and High, K.A. (2011). Therapeutic in vivo gene transfer for genetic disease using AAV: progress and challenges. *Nat. Rev. Genet.* 12, 341–355.
- Zhang, W., Hagedorn, C., Schulz, E., Lipps, H.J., and Ehrhardt, A. (2014). Viral hybrid-vectors for delivery of autonomous replicons. *Curr. Gene Ther.* 14, 10–23.
- Cooney, A.L., McCray, P.B., Jr., and Sinn, P.L. (2018). Cystic Fibrosis Gene Therapy: Looking Back, Looking Forward. *Genes (Basel)* 9, E538.
- Wang, L., Zheng, W., Liu, S., Li, B., and Jiang, X. (2019). Delivery of CRISPR/Cas9 by novel strategies for gene therapy. *ChemBioChem* 20, 634–643.
- Degtiar, I. (2017). A review of international coverage and pricing strategies for personalized medicine and orphan drugs. *Health Policy* 121, 1240–1248.
- Li, B., He, X., Jia, W., and Li, H. (2017). Novel Applications of Metabolomics in Personalized Medicine: A Mini-Review. *Molecules* 22, E1173.
- Larionov, V., Kouprina, N., Graves, J., Chen, X.N., Korenberg, J.R., and Resnick, M.A. (1996). Specific cloning of human DNA as yeast artificial chromosomes by transformation-associated recombination. *Proc. Natl. Acad. Sci. USA* 93, 491–496.
- Larionov, V., Kouprina, N., Graves, J., and Resnick, M.A. (1996). Highly selective isolation of human DNAs from rodent-human hybrid cells as circular yeast artificial chromosomes by transformation-associated recombination cloning. *Proc. Natl. Acad. Sci. USA* 93, 13925–13930.
- Larionov, V., Kouprina, N., Solomon, G., Barrett, J.C., and Resnick, M.A. (1997). Direct isolation of human BRCA2 gene by transformation-associated recombination in yeast. *Proc. Natl. Acad. Sci. USA* 94, 7384–7387.
- Kouprina, N., Graves, J., Cancilla, M.R., Resnick, M.A., and Larionov, V. (1997). Specific isolation of human rDNA genes by TAR cloning. *Gene* 197, 269–276.
- Kouprina, N., Annab, L., Graves, J., Afshari, C., Barrett, J.C., Resnick, M.A., and Larionov, V. (1998). Functional copies of a human gene can be directly isolated by transformation-associated recombination cloning with a small 3' end target sequence. *Proc. Natl. Acad. Sci. USA* 95, 4469–4474.
- Cancilla, M.R., Tainton, K.M., Barry, A.E., Larionov, V., Kouprina, N., Resnick, M.A., Sart, D.D., and Choo, K.H. (1998). Direct cloning of human 10q25 neocentromere DNA using transformation-associated recombination (TAR) in yeast. *Genomics* 47, 399–404.
- Cancilla, M.R., Graves, J., Matesic, L.E., Reeves, R.H., Tainton, K.M., Choo, K.H., Resnick, M.A., Larionov, V.L., and Kouprina, N.Y. (1998). Rapid cloning of mouse DNA as yeast artificial chromosomes by transformation-associated recombination (TAR). *Mamm. Genome* 9, 157–159.
- Kouprina, N., Kawamoto, K., Barrett, J.C., Larionov, V., and Koi, M. (1998). Rescue of targeted regions of mammalian chromosomes by in vivo recombination in yeast. *Genome Res.* 8, 666–672.
- Kouprina, N., Campbell, M., Graves, J., Campbell, E., Meincke, L., Tesmer, J., Grady, D.L., Doggett, N.A., Moyzis, R.K., Deaven, L.L., and Larionov, V. (1998). Construction of human chromosome 16- and 5-specific circular YAC/BAC libraries by in vivo recombination in yeast (TAR cloning). *Genomics* 53, 21–28.
- Annab, L.A., Kouprina, N., Solomon, G., Cable, P.L., Hill, D.E., Barrett, J.C., Larionov, V., and Afshari, C.A. (2000). Isolation of a functional copy of the human BRCA1 gene by transformation-associated recombination in yeast. *Gene* 250, 201–208.
- Kim, J., Noskov, V.N., Lu, X., Bergmann, A., Ren, X., Warth, T., Richardson, P., Kouprina, N., and Stubbs, L. (2000). Discovery of a novel, paternally expressed ubiquitin-specific processing protease gene through comparative analysis of an imprinted region of mouse chromosome 7 and human chromosome 19q13.4. *Genome Res.* 10, 1138–1147.
- Humble, M.C., Kouprina, N., Noskov, V.N., Graves, J., Garner, E., Tennant, R.W., Resnick, M.A., Larionov, V., and Cannon, R.E. (2000). Radial transformation-associated recombination cloning from the mouse genome: isolation of Tg.AC transgene with flanking DNAs. *Genomics* 70, 292–299.
- Zeng, C., Kouprina, N., Zhu, B., Cairo, A., Hoek, M., Cross, G., Osoegawa, K., Larionov, V., and de Jong, P. (2001). Large-insert BAC/YAC libraries for selective re-isolation of genomic regions by homologous recombination in yeast. *Genomics* 77, 27–34.
- Leem, S.-H., Londoño-Vallejo, J.A., Kim, J.-H., Bui, H., Tubacher, E., Solomon, G., Park, J.-E., Horikawa, I., Kouprina, N., Barrett, J.C., and Larionov, V. (2002). The human telomerase gene: complete genomic sequence and analysis of tandem repeat polymorphisms in intronic regions. *Oncogene* 21, 769–777.
- Nihei, N., Kouprina, N., Larionov, V., Oshima, J., Martin, G.M., Ichikawa, T., and Barrett, J.C. (2002). Functional evidence for a metastasis suppressor gene for rat prostate cancer within a 60-kilobase region on human chromosome 8p21-p12. *Cancer Res.* 62, 367–370.
- Kouprina, N., Noskov, V.N., Koriabine, M., Leem, S.H., and Larionov, V. (2004). Exploring transformation-associated recombination cloning for selective isolation of genomic regions. *Methods Mol. Biol.* 255, 69–89.
- Kouprina, N., and Larionov, V. (2006). TAR cloning: insights into gene function, structural variants, long-range haplotypes and genome evolution. *Nat. Rev. Genet.* 7, 805–812.
- Kouprina, N., and Larionov, V. (2008). Selective isolation of genomic loci from complex genomes by transformation-associated recombination cloning in the yeast *Saccharomyces cerevisiae*. *Nat. Protoc.* 3, 371–377.
- Kouprina, N., and Larionov, V. (2016). Transformation-associated recombination (TAR) cloning for genomics studies and synthetic biology. *Chromosoma* 125, 621–632.
- Lee, N.C.O., Larionov, V., and Kouprina, N. (2015). Highly efficient CRISPR/Cas9-mediated TAR cloning of genes and chromosomal loci from complex genomes in yeast. *Nucleic Acids Res.* 43, e55.
- Kouprina, N., Mullokandov, M., Rogozin, I.B., Collins, N.K., Solomon, G., Risinger, J.L., Koonin, E.V., Barrett, J.C., and Larionov, V. (2004). The SPANX gene family of cancer/testis-specific antigens: rapid evolution and amplification in African great apes and hominids. *Proc. Natl. Acad. Sci. USA* 101, 3077–3082.
- Kouprina, N., Pavlicek, A., Mochida, G.H., Solomon, G., Gersch, W., Yoon, Y.H., Collura, R., Ruvolo, M., Barrett, J.C., Woods, C.G., et al. (2004). Accelerated evolution of the ASPM gene controlling brain size begins prior to human brain expansion. *PLoS Biol.* 2, E126.
- Pavlicek, A., Noskov, V.N., Kouprina, N., Barrett, J.C., Jurka, J., and Larionov, V. (2004). Evolution of the tumor suppressor BRCA1 locus in primates: implications for cancer predisposition. *Hum. Mol. Genet.* 13, 2737–2751.
- Kouprina, N., Noskov, V.N., Pavlicek, A., Collins, N.K., Schoppee Bortz, P.D., Ottolenghi, C., Loukinov, D., Goldsmith, P., Risinger, J.L., Kim, J.H., et al. (2007). Evolutionary Diversification of SPANX-N Sperm Protein Gene Structure and Expression. *PLoS One* 2, e359.
- Kim, J.H., Leem, S.H., Sunwoo, Y., and Kouprina, N. (2003). Separation of long-range human TERT gene haplotypes by transformation-associated recombination cloning in yeast. *Oncogene* 22, 2452–2456.
- Kouprina, N., Pavlicek, A., Noskov, V.N., Solomon, G., Otstot, J., Isaacs, W., Carpten, J.D., Trent, J.M., Schleutker, J., Barrett, J.C., et al. (2005). Dynamic structure of the SPANX gene cluster mapped to the prostate cancer susceptibility locus HPCX at Xq27. *Genome Res.* 15, 1477–1486.



35. Kouprina, N., Noskov, V.N., Solomon, G., Otsot, J., Isaacs, W., Xu, J., Schleutker, J., and Larionov, V. (2007). Mutational analysis of SPANX genes in families with X-linked prostate cancer. *Prostate* 67, 820–828.
36. Kouprina, N., Lee, N.C., Pavlicek, A., Samoshkin, A., Kim, J.H., Lee, H.S., Varma, S., Reinhold, W.C., Otsot, J., Solomon, G., et al. (2012). Exclusion of the 750-kb genetically unstable region at Xq27 as a candidate locus for prostate malignancy in HPCX1-linked families. *Genes Chromosomes Cancer* 51, 933–948.
37. Kouprina, N., Liskovych, M., Lee, N.C.O., Noskov, V.N., Waterfall, J.J., Walker, R.L., Meltzer, P.S., Topol, E.J., and Larionov, V. (2018). Analysis of the 9p21.3 sequence associated with coronary artery disease reveals a tendency for duplication in a CAD patient. *Oncotarget* 9, 15275–15291.
38. Noskov, V.N., Leem, S.H., Solomon, G., Mullokandov, M., Chae, J.Y., Yoon, Y.H., Shin, Y.S., Kouprina, N., and Larionov, V. (2003). A novel strategy for analysis of gene homologues and segmental genome duplications. *J. Mol. Evol.* 56, 702–710.
39. Leem, S.H., Kouprina, N., Grimwood, J., Kim, J.H., Mullokandov, M., Yoon, Y.H., Chae, J.Y., Morgan, J., Lucas, S., Richardson, P., et al. (2004). Closing the gaps on human chromosome 19 revealed genes with a high density of repetitive tandemly arrayed elements. *Genome Res.* 14, 239–246.
40. Kouprina, N., Leem, S.H., Solomon, G., Ly, A., Koriabine, M., Otsot, J., Pak, E., Dutra, A., Zhao, S., Barrett, J.C., and Larionov, V. (2003). Segments missing from the draft human genome sequence can be isolated by transformation-associated recombination cloning in yeast. *EMBO Rep.* 4, 257–262.
41. Grimwood, J., Gordon, L.A., Olsen, A., Terry, A., Schmutz, J., Lamerding, J., Hellsten, U., Goodstein, D., Couronne, O., Tran-Gyamfi, M., et al. (2004). The DNA sequence and biology of human chromosome 19. *Nature* 428, 529–535.
42. Kim, J.H., Dilthey, A.T., Nagaraja, R., Lee, H.S., Koren, S., Dudekula, D., Wood Iii, W.H., Piao, Y., Ogurtsov, A.Y., Utani, K., et al. (2018). Variation in human chromosome 21 ribosomal RNA genes characterized by TAR cloning and long-read sequencing. *Nucleic Acids Res.* 46, 6712–6725.
43. Nowakowska, B. (2017). Clinical interpretation of copy number variants in the human genome. *J. Appl. Genet.* 58, 449–457.
44. Lew, A.R., Keller Mayer, T.R., Sule, B.P., and Szigeti, K. (2018). Copy Number Variations in Adult-onset Neuropsychiatric Diseases. *Curr. Genomics* 19, 420–430.
45. Salpietro, V., Manole, A., Efthymiou, S., and Houlden, H. (2018). A Review of Copy Number Variants in Inherited Neuropathies. *Curr. Genomics* 19, 412–419.
46. Morin, S.J., Eccles, J., Iturriaga, A., and Zimmerman, R.S. (2017). Translocations, inversions and other chromosome rearrangements. *Fertil. Steril.* 107, 19–26.
47. Puig, M., Casillas, S., Villatoro, S., and Cáceres, M. (2015). Human inversions and their functional consequences. *Brief. Funct. Genomics* 14, 369–379.
48. Kuo, C.Y. (2018). Advances in site-specific gene editing for primary immune deficiencies. *Curr. Opin. Allergy Clin. Immunol.* 18, 453–458.
49. Jenks, S. (2017). New Gene-Editing Tool Slowly Advances Into the Clinic. *J. Natl. Cancer Inst.* 109.
50. Baliou, S., Adamaki, M., Kyriakopoulos, A.M., Spandidos, D.A., Panayiotidis, M., Christodoulou, I., and Zoumpourlis, V. (2018). CRISPR therapeutic tools for complex genetic disorders and cancer (Review). *Int. J. Oncol.* 53, 443–468.
51. Jain, M., Koren, S., Miga, K.H., Quick, J., Rand, A.C., Sasani, T.A., Tyson, J.R., Beggs, A.D., Dilthey, A.T., Fiddes, I.T., et al. (2018). Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nat. Biotechnol.* 36, 338–345.
52. Gabrieli, T., Sharim, H., Fridman, D., Arbib, N., Michaeli, Y., and Ebenstein, Y. (2018). Selective nanopore sequencing of human BRCA1 by Cas9-assisted targeting of chromosome segments (CATCH). *Nucleic Acids Res.* 46, e87.
53. Baffoe-Bonnie, A.B., Smith, J.R., Stephan, D.A., Schleutker, J., Carpten, J.D., Kainu, T., Gillanders, E.M., Matikainen, M., Teslovich, T.M., Tammela, T., et al. (2005). A major locus for hereditary prostate cancer in Finland: localization by linkage disequilibrium of a haplotype in the HPCX region. *Hum. Genet.* 117, 307–316.
54. Wang, Z., McGlynn, K.A., Rajpert-De Meyts, E., Bishop, D.T., Chung, C.C., Dalgaard, M.D., Greene, M.H., Gupta, R., Grotmol, T., Haugen, T.B., et al.; Testicular Cancer Consortium (2017). Meta-analysis of five genome-wide association studies identifies multiple new loci associated with testicular germ cell tumor. *Nat. Genet.* 49, 1141–1147.
55. Huc-Chabrolle, M., Charon, C., Guilmatre, A., Vourc'h, P., Tripi, G., Barthez, M.A., Sizaret, E., Thepault, R.A., Le Gallic, S., Hager, J., et al. (2013). Xq27 FRAXA locus is a strong candidate for dyslexia: evidence from a genome-wide scan in French families. *Behav. Genet.* 43, 132–140.
56. Chernoff, Y.O., Newnam, G.P., and Liebman, S.W. (1996). The translational function of nucleotide C1054 in the small subunit rRNA is conserved throughout evolution: genetic evidence in yeast. *Proc. Natl. Acad. Sci. USA* 93, 2517–2522.
57. Porokhovnik, L.N., and Lyapunova, N.A. (2019). Dosage effects of human ribosomal genes (rDNA) in health and disease. *Chromosome Res.* 27, 5–17.
58. Lyapunova, N.A., Porokhovnik, L.N., Kosyakova, N.V., Mandron, I.A., and Tsvetkova, T.G. (2017). Effects of the copy number of ribosomal genes (genes for rRNA) on viability of subjects with chromosomal abnormalities. *Gene* 611, 47–53.
59. Ding, Q., Markesbery, W.R., Chen, Q., Li, F., and Keller, J.N. (2005). Ribosome dysfunction is an early event in Alzheimer's disease. *J. Neurosci.* 25, 9171–9175.
60. Chestkov, I.V., Jestkova, E.M., Ershova, E.S., Golimbet, V.E., Lezheiko, T.V., Kolesina, N.Y., Porokhovnik, L.N., Lyapunova, N.A., Izhevskaya, V.L., Kutsev, S.I., et al. (2018). Abundance of ribosomal RNA gene copies in the genomes of schizophrenia patients. *Schizophr. Res.*, Published online January 11, 2018. <https://doi.org/10.1016/j.schres.2018.01.001>.
61. Kouprina, N., Earnshaw, W.C., Masumoto, H., and Larionov, V. (2013). A new generation of human artificial chromosomes for functional genomics and gene therapy. *Cell. Mol. Life Sci.* 70, 1135–1148.
62. Kouprina, N., Tomilin, A.N., Masumoto, H., Earnshaw, W.C., and Larionov, V. (2014). Human artificial chromosome-based gene delivery vectors for biomedicine and biotechnology. *Expert Opin. Drug Deliv.* 11, 517–535.
63. Kazuki, Y., and Oshimura, M. (2011). Human artificial chromosomes for gene delivery and the development of animal models. *Mol. Ther.* 19, 1591–1601.
64. Oshimura, M., Uno, N., Kazuki, Y., Katoh, M., and Inoue, T. (2015). A pathway from chromosome transfer to engineering resulting in human and mouse artificial chromosomes for a variety of applications to bio-medical challenges. *Chromosome Res.* 23, 111–133.
65. Kouprina, N., Petrov, N., Molina, O., Liskovych, M., Pesenti, E., Ohzeki, J.L., Masumoto, H., Earnshaw, W.C., and Larionov, V. (2018). Human Artificial Chromosome with Regulated Centromere: A Tool for Genome and Cancer Studies. *ACS Synth. Biol.* 7, 1974–1989.
66. Katona, R.L. (2015). De novo formed satellite DNA-based mammalian artificial chromosomes and their possible applications. *Chromosome Res.* 23, 143–157.
67. Moralli, D., and Monaco, Z.L. (2015). Developing de novo human artificial chromosomes in embryonic stem cells using HSV-1 amplicon technology. *Chromosome Res.* 23, 105–110.
68. Ohzeki, J., Larionov, V., Earnshaw, W.C., and Masumoto, H. (2015). Genetic and epigenetic regulation of centromeres: a look at HAC formation. *Chromosome Res.* 23, 87–103.
69. Molina, O., Kouprina, N., Masumoto, H., Larionov, V., and Earnshaw, W.C. (2017). Using human artificial chromosomes to study centromere assembly and function. *Chromosoma* 126, 559–575.
70. Ebersole, T., Okamoto, Y., Noskov, V.N., Kouprina, N., Kim, J.-H., Leem, S.-H., Barrett, J.C., Masumoto, H., and Larionov, V. (2005). Rapid generation of long synthetic tandem repeats and its application for analysis in human artificial chromosome formation. *Nucleic Acids Res.* 33, e130.
71. Noskov, V.N., Lee, N.C., Larionov, V., and Kouprina, N. (2011). Rapid generation of long tandem DNA repeat arrays by homologous recombination in yeast to study their function in mammalian genomes. *Biol. Proced. Online* 13, 8.
72. Nakano, M., Cardinale, S., Noskov, V.N., Gassmann, R., Vagnarelli, P., Kandels-Lewis, S., Larionov, V., Earnshaw, W.C., and Masumoto, H. (2008). Inactivation of a human kinetochore by specific targeting of chromatin modifiers. *Dev. Cell* 14, 507–522.
73. Kim, J.H., Kononenko, A., Erliandri, I., Kim, T.A., Nakano, M., Iida, Y., Barrett, J.C., Oshimura, M., Masumoto, H., Earnshaw, W.C., et al. (2011). Human artificial chromosome (HAC) vector with a conditional centromere for correction of genetic deficiencies in human cells. *Proc. Natl. Acad. Sci. USA* 108, 20048–20053.



74. Kononenko, A.V., Bansal, R., Lee, N.C., Grimes, B.R., Masumoto, H., Earnshaw, W.C., Larionov, V., and Kouprina, N. (2014). A portable BRCA1-HAC (human artificial chromosome) module for analysis of BRCA1 tumor suppressor function. *Nucleic Acids Res.* **42**.
75. Tedesco, F.S., Hoshiya, H., D'Antona, G., Gerli, M.F., Messina, G., Antonini, S., Tonlorenzi, R., Benedetti, S., Berghella, L., Torrente, Y., et al. (2011). Stem cell-mediated transfer of a human artificial chromosome ameliorates muscular dystrophy. *Sci. Transl. Med.* **3**, 96ra78.
76. Benedetti, S., Uno, N., Hoshiya, H., Ragazzi, M., Ferrari, G., Kazuki, Y., Moyle, L.A., Tonlorenzi, R., Lombardo, A., Chaouch, S., et al. (2018). Reversible immortalisation enables genetic correction of human muscle progenitors and engineering of next-generation human artificial chromosomes for Duchenne muscular dystrophy. *EMBO Mol. Med.* **10**, 254–275.
77. Kazuki, Y., Hiratsuka, M., Takiguchi, M., Osaki, M., Kajitani, N., Hoshiya, H., Hiratsuka, K., Yoshino, T., Kazuki, K., Ishihara, C., et al. (2010). Complete genetic correction of ips cells from Duchenne muscular dystrophy. *Mol. Ther.* **18**, 386–393.
78. Sinenko, S.A., Skvortsova, E.V., Liskovych, M.A., Ponomartsev, S.V., Kuzmin, A.A., Khudiakov, A.A., Malashicheva, A.B., Alenina, N., Larionov, V., Kouprina, N., and Tomilin, A.N. (2018). Transfer of synthetic human chromosome into human induced pluripotent stem cells for biomedical applications. *Cells* **7**, E261.
79. Liskovych, M., Ponomartsev, S., Popova, E., Bader, M., Kouprina, N., Larionov, V., Alenina, N., and Tomilin, A. (2015). Stable maintenance of de novo assembled human artificial chromosomes in embryonic stem cells and their differentiated progeny in mice. *Cell Cycle* **14**, 1268–1273.
80. Kobayashi, K., Abe, C., Endo, M., Kazuki, Y., Oshimura, M., and Chiba, K. (2017). Gender Difference of Hepatic and Intestinal CYP3A4 in CYP3A Humanized Mice Generated by a Human Chromosome-engineering Technique. *Drug Metab. Lett.* **11**, 60–67.
81. Watanabe, Y., Kazuki, Y., Kazuki, K., Ebiki, M., Nakanishi, M., Nakamura, K., Yoshida Yamakawa, M., Hosokawa, H., Ohbayashi, T., Oshimura, M., and Nakashima, K. (2015). Use of a Human Artificial Chromosome for Delivering Trophic Factors in a Rodent Model of Amyotrophic Lateral Sclerosis. *Mol. Ther. Nucleic Acids* **4**, e253.
82. Kazuki, Y., Kobayashi, K., Aueviriyavit, S., Oshima, T., Kuroiwa, Y., Tsukazaki, Y., Senda, N., Kawakami, H., Ohtsuki, S., Abe, S., et al. (2013). Trans-chromosomal mice containing a human CYP3A cluster for prediction of xenobiotic metabolism in humans. *Hum. Mol. Genet.* **22**, 578–592.
83. Lee, N.C., Kononenko, A.V., Lee, H.S., Tolkunova, E.N., Liskovych, M.A., Masumoto, H., Earnshaw, W.C., Tomilin, A.N., Larionov, V., and Kouprina, N. (2013). Protecting a transgene expression from the HAC-based vector by different chromatin insulators. *Cell. Mol. Life Sci.* **70**, 3723–3737.
84. Lee, N.C.O., Kim, J.H., Petrov, N.S., Lee, H.S., Masumoto, H., Earnshaw, W.C., Larionov, V., and Kouprina, N. (2018). Method to Assemble Genomic DNA Fragments or Genes on Human Artificial Chromosome with Regulated Kinetochores Using a Multi-Integrase System. *ACS Synth. Biol.* **7**, 63–74.
85. Honma, K., Abe, S., Endo, T., Uno, N., Oshimura, M., Ohbayashi, T., and Kazuki, Y. (2018). Development of a multiple-gene-loading method by combining multi-integration system-equipped mouse artificial chromosome vector and CRISPR-Cas9. *PLoS ONE* **13**, e0193642.
86. Tóth, A., Fodor, K., Praznovszky, T., Tubak, V., Udvardy, A., Hadlaczky, G., and Katona, R.L. (2014). Novel method to load multiple genes onto a mammalian artificial chromosome. *PLoS ONE* **9**, e85565.
87. Ross, A.C., Gulland, L.E., Dorrestein, P.C., and Moore, B.S. (2015). Targeted capture and heterologous expression of the Pseudoalteromonas alterochromide gene cluster in *Escherichia coli* represents a promising natural product exploratory platform. *ACS Synth. Biol.* **4**, 414–420.
88. Cano-Prieto, C., García-Salcedo, R., Sánchez-Hidalgo, M., Braña, A.F., Fiedler, H.P., Méndez, C., Salas, J.A., and Olano, C. (2015). Genome mining of streptomyces sp. Tü 6176: characterization of the natakazole biosynthesis pathway. *ChemBioChem* **16**, 1461–1473.
89. Yamanaka, K., Reynolds, K.A., Kersten, R.D., Ryan, K.S., Gonzalez, D.J., Nizet, V., Dorrestein, P.C., and Moore, B.S. (2014). Direct cloning and refactoring of a silent lipopeptide biosynthetic gene cluster yields the antibiotic taromycin A. *Proc. Natl. Acad. Sci. USA* **111**, 1957–1962.
90. Bonet, B., Teufel, R., Crüsemann, M., Ziemert, N., and Moore, B.S. (2015). Direct capture and heterologous expression of *Salinispora* natural product genes for the biosynthesis of enterocin. *J. Nat. Prod.* **78**, 539–542.
91. Li, Y., Li, Z., Yamanaka, K., Xu, Y., Zhang, W., Vlamakis, H., Kolter, R., Moore, B.S., and Qian, P.Y. (2015). Directed natural product biosynthesis gene cluster capture and expression in the model bacterium *Bacillus subtilis*. *Sci. Rep.* **5**, 9383.
92. García-Salcedo, R., Álvarez-Álvarez, R., Olano, C., Cañedo, L., Braña, A.F., Méndez, C., de la Calle, F., and Salas, J.A. (2018). Characterization of the Jomthonic Acid Biosynthesis Pathway and Isolation of Novel Analogues in *Streptomyces caniferus* GUA-06-05-006A. *Mar. Drugs* **16**, E259.
93. Shao, Z., Zhao, H., and Zhao, H. (2009). DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways. *Nucleic Acids Res.* **37**, e16.
94. Ongley, S.E., Bian, X., Neilan, B.A., and Müller, R. (2013). Recent advances in the heterologous expression of microbial natural product biosynthetic pathways. *Nat. Prod. Rep.* **30**, 1121–1138.
95. Hu, Y., Nan, F., Maina, S.W., Guo, J., Wu, S., and Xin, Z. (2018). Clone of plipastatin biosynthetic gene cluster by transformation-associated recombination technique and high efficient expression in model organism *Bacillus subtilis*. *J. Biotechnol.* **288**, 1–8.
96. Bilyk, O., Sekurova, O.N., Zotchev, S.B., and Luzhetskyy, A. (2016). Cloning and Heterologous Expression of the Grecoacycline Biosynthetic Gene Cluster. *PLoS ONE* **11**, e0158682.
97. Wu, N., Huang, H., Min, T., and Hu, H. (2017). TAR cloning and integrated over-expression of 6-demethylchlorotetracycline biosynthetic gene cluster in *Streptomyces aureofaciens*. *Acta Biochim. Biophys. Sin. (Shanghai)* **49**, 1129–1134.
98. Tang, X., Li, J., Millán-Aguñaga, N., Zhang, J.J., O'Neill, E.C., Ugalde, J.A., Jensen, P.R., Mantovani, S.M., and Moore, B.S. (2015). Identification of Thiotetronic Acid Antibiotic Biosynthetic Pathways by Target-directed Genome Mining. *ACS Chem. Biol.* **10**, 2841–2849.
99. Novakova, R., Núñez, L.E., Homerova, D., Knirschova, R., Feckova, L., Rezuchova, B., Sevcikova, B., Menéndez, N., Moris, F., Cortés, J., and Kormanec, J. (2018). Increased heterologous production of the antitumoral polyketide mithramycin A by engineered *Streptomyces lividans* TK24 strains. *Appl. Microbiol. Biotechnol.* **102**, 857–869.
100. Meng, J., Feng, R., Zheng, G., Ge, M., Mast, Y., Wohlleben, W., Gao, J., Jiang, W., and Lu, Y. (2017). Improvement of pristinamycin I (PI) production in *Streptomyces pristinaespiralis* by metabolic engineering approaches. *Synth Syst Biotechnol* **2**, 130–136.
101. Alberti, F., Leng, D.J., Wilkening, I., Song, L., Tosin, M., and Corre, C. (2018). Triggering the expression of a silent gene cluster from genetically intractable bacteria results in scleric acid discovery. *Chem. Sci. (Camb.)* **10**, 453–463.
102. Newman, D.J., and Cragg, G.M. (2016). Natural Products as Sources of New Drugs from 1981 to 2014. *J. Nat. Prod.* **79**, 629–661.
103. Noskov, V.N., Kouprina, N., Leem, S.H., Ouspenski, I., Barrett, J.C., and Larionov, V. (2003). A general cloning system to selectively isolate any eukaryotic or prokaryotic genomic region in yeast. *BMC Genomics* **4**, 16.
104. Huo, L., Hug, J.J., Fu, C., Bian, X., Zhang, Y., and Müller, R. (2019). Heterologous expression of bacterial natural product biosynthetic pathways. *Nat. Prod. Rep.* , Published online January 8, 2019. <https://doi.org/10.1039/c8np00091c>.
105. Benders, G.A., Noskov, V.N., Denisova, E.A., Lartigue, C., Gibson, D.G., Assad-Garcia, N., Chuang, R.Y., Carrera, W., Moodie, M., Algire, M.A., et al. (2010). Cloning whole bacterial genomes in yeast. *Nucleic Acids Res.* **38**, 2558–2569.
106. Zhou, J., Wu, R., Xue, X., and Qin, Z. (2016). CasHRA (Cas9-facilitated Homologous Recombination Assembly) method of constructing megabase-sized DNA. *Nucleic Acids Res.* **44**, e124.
107. Tagwerker, C., Dupont, C.L., Karas, B.J., Ma, L., Chuang, R.Y., Benders, G.A., Ramon, A., Novotny, M., Montague, M.G., Venepally, P., et al. (2012). Sequence analysis of a complete 1.66 Mb *Prochlorococcus marinus* MED4 genome cloned in yeast. *Nucleic Acids Res.* **40**, 10375–10383.
108. Satoh, D., Abe, S., Kobayashi, K., Nakajima, Y., Oshimura, M., and Kazuki, Y. (2018). Human and mouse artificial chromosome technologies for studies



- of pharmacokinetics and toxicokinetics. *Drug Metab. Pharmacokinet.* 33, 17–30.
109. Fachinetti, D., Logsdon, G.A., Abdullah, A., Selzer, E.B., Cleveland, D.W., and Black, B.E. (2017). CENP-A modifications on Ser68 and Lys124 are dispensable for establishment, maintenance, and long-term function of human centromeres. *Dev. Cell* 40, 104–113.
110. Logsdon, G.A., and Black, B.E. (2018). The new normal of structure/function studies in the era of CRISPR/Cas9. *Biochem. J.* 475, 1635–1642.
111. Noskov, V.N., Koriabine, M., Solomon, G., Randolph, M., Barrett, J.C., Leem, S.-H., Stubbs, L., Kouprina, N., and Larionov, V. (2001). Defining the minimal length of sequence homology required for selective gene isolation by TAR cloning. *Nucleic Acids Res.* 29, E32.
112. Liskovych, M., Lee, N.C.O., Larionov, V., and Kouprina, N. (2016). Moving toward a higher efficiency of microcell-mediated chromosome transfer. *Mol. Ther. Methods Clin. Dev.* 3, 16043.
113. Suzuki, T., Kazuki, Y., Oshimura, M., and Hara, T. (2016). Highly Efficient Transfer of Chromosomes to a Broad Range of Target Cells Using Chinese Hamster Ovary Cells Expressing Murine Leukemia Virus-Derived Envelope Proteins. *PLoS ONE* 11, e0157187.
114. Hasegawa, Y., Ikeno, M., Suzuki, N., Nakayama, M., and Ohara, O. (2018). Improving the efficiency of gene insertion in a human artificial chromosome vector and its transfer in human-induced pluripotent stem cells. *Biol. Methods Protoc.* 3, bpy013.
115. Park, I.H. (2010). DYS-HAC-iPS cells: the combination of gene and cell therapy to treat duchenne muscular dystrophy. *Mol. Ther.* 18, 238–240.
116. Yakura, Y., Ishihara, C., Kurosaki, H., Kazuki, Y., Komatsu, N., Okada, Y., Doi, T., Takeya, H., and Oshimura, M. (2013). An induced pluripotent stem cell-mediated and integration-free factor VIII expression system. *Biochem. Biophys. Res. Commun.* 431, 336–341.