



Rapid assembly of functional modules for generating human artificial chromosome constructs compatible with epigenetic centromere seeding

Gabriel J. Birchak · Daniel G. Gibson · Praveen Kumar Allu ·
Prakriti Kashyap · John I. Glass · Ben E. Black

Received: 23 September 2025 / Revised: 11 November 2025 / Accepted: 13 November 2025 / Published online: 1 December 2025
© The Author(s) 2025

Abstract The ongoing development of human artificial chromosomes (HACs) will permit investigation into essential centromere processes and the means to deliver large genetic cargoes to target cells. Starting with large (~750 kb) yeast artificial chromosome (YAC)-based constructs limits the rampant multimerization that has complicated many prior types of HACs. Large YAC construction is accomplished using transformation-associated recombination (TAR) strategies that can become unwieldy when several functional modules are to be incorporated and tested. To address this issue, we developed an approach where modules are built using high-fidelity in vitro assembly strategies in a bacterial artificial chromosome (BAC) format. Then, the assembled

modules are transferred in a simplified TAR step into a recipient YAC harboring the prokaryotic “stuffer” DNA that comprises a large portion of the final HAC construct. This approach is highly efficient with two-thirds of all screened yeast clones harboring the correct TAR product. Further, whole-genome Oxford Nanopore Technologies (ONT) sequencing/alignments, de novo assembly of the final YAC using a single ONT sequencing run, and close inspection of highly repetitive regions are all streamlined to rapidly validate clones that match the design. The fully sequenced, verified strain harboring a multi-module construct was then fused to human cells, where it efficiently formed functional HACs upon initial seeding with CENP-A-containing nucleosomes. We envision that the rapid assembly steps will be useful to quickly incorporate different functional modules, including diverse genetic cargoes, to engineer HACs with specific design features.

Responsible Editor: Bungo Akiyoshi

Gabriel J. Birchak and Daniel G. Gibson contributed equally to this work.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10577-025-09788-w>.

G. J. Birchak · P. K. Allu · P. Kashyap · B. E. Black (✉)
Department of Biochemistry & Biophysics, Perelman
School of Medicine, University of Pennsylvania,
Philadelphia, PA 19104, USA
e-mail: blackbe@pennmedicine.upenn.edu

G. J. Birchak · P. K. Allu · P. Kashyap · B. E. Black
Penn Center for Genome Integrity, University
of Pennsylvania, Philadelphia, PA 19104, USA

G. J. Birchak · P. K. Allu · P. Kashyap · B. E. Black
Epigenetics Institute, University of Pennsylvania,
Philadelphia, PA 19104, USA

G. J. Birchak · B. E. Black
Graduate Program in Cell & Molecular Biology,
University of Pennsylvania, Philadelphia, PA 19104, USA

D. G. Gibson · J. I. Glass
J. Craig Venter Institute, La Jolla, CA 92037, USA

Keywords Human artificial chromosomes (HACs) · Oxford nanopore technologies (ONT) sequencing · Genomics · Chromosome biology · Centromere biology · Epigenetics · Transformation-associated recombination (TAR) cloning

Introduction

Human artificial chromosomes (HACs) that form de novo in recipient human cells have typically been derived from DNA introduced by bacterial artificial chromosomes (BACs) or yeast artificial chromosomes (YACs) that then acquires the ability to replicate in S-phase and subsequently segregate like natural chromosomes to daughter cells at cell division (Harrington et al. 1997; Ikeno et al. 1998; Okada et al. 2007; Logsdon et al. 2019). Chromosome segregation is mediated by the centromere (Kixmoeller et al. 2020; Musacchio and Desai 2017), which is defined by chromatin harboring the histone H3 variant, CENP-A (Earnshaw and Rothfield 1985; Palmer et al. 1987). CENP-A nucleosomes recruit a 16-protein complex, the constitutive centromere-associated network (CCAN) (Amano et al. 2009; Foltz et al. 2006; Hori et al. 2008; Okada et al. 2006), that is present throughout the cell cycle and provides the foundation for the mitotic kinetochore that connects to spindle microtubules. Rare, spontaneous acquisition of functional centromeric chromatin can be used to identify DNA sequence features in human α -satellite DNA that contribute to natural chromosome behavior (Hayden et al. 2013; Maloney et al. 2012; Ohzeki et al. 2002; Schueler et al. 2001). Efficient centromere formation on α -satellite or even non-repetitive DNA lacking sequence homology to α -satellite DNA can be achieved by experimentally directing CENP-A nucleosome assembly (Barnhart et al. 2011; Chen et al. 2014; Hori et al. 2012; Mendiburo et al. 2011). The same approach has facilitated more recent efforts in the efficient formation of HACs (Logsdon et al. 2019).

The majority of de novo HAC formation approaches since their inception involved initial DNA constructs in the 40–200 kb range that invariably were rearranged (with themselves or with host cell DNA) and then multimerized many times to final sizes in the megabase range (Ebersole et al. 2000; Harrington et al. 1997; Ikeno et al. 1998; Kouprina et al. 2012;

Logsdon et al. 2019; Rudd et al. 2003). These rearrangements and amplifications are not compatible with control over the genetic material carried by the HACs. Using a larger initial YAC construct (~750 kb)—*sufficient for chromatin spanning from the kinetochore to the inner centromere where sister chromatin cohesion is maintained until the moment of their separation at the onset of anaphase*—along with epigenetic centromere seeding and an improved cell delivery strategy successfully leads to efficient HAC formation without rampant multimerization (Gambogi et al. 2024). Outfitting the ~750 kb YAC involved a cloning strategy that required the simultaneous recombination of several different DNA fragments in yeast by transformation-associated recombination (TAR) (Gambogi et al. 2024). Simultaneous assembly of multiple DNA fragments substantially reduces TAR efficiency (Gibson et al. 2008a, 2008b), however, and is likely to do so to such an extent with some designs as to preclude isolation of the desired YAC.

Here, we report an approach that combines modules in high-efficiency steps that precede a TAR step to introduce designed HAC modules into a general YAC encoding the prokaryotic “stuffer” DNA. The initial steps involve rapid cycles of standard isothermal assembly, isolation in *E. coli*, and DNA sequence validation. After multiple modules are assembled in a BAC, it is introduced via a simplified high-efficiency TAR reaction. Further, the YAC built from component modules in this fashion efficiently forms functional HACs in recipient human cells by epigenetic centromere seeding.

Results and discussion

Combining functional modules prior to TAR-mediated assembly of a HAC construct

We redesigned the strategy (Fig. 1) to assemble a functional HAC construct with the same basic features of one that we recently reported, termed YAC-*Mm*-4q21^{LacO} (Gambogi et al. 2024). As in the initial effort, the starting point is a YAC harboring the *M. mycoides* ~550 kb synthetic prokaryotic genome (JCVI-syn3B; (Bittencourt et al. 2024)) that will serve as stuffer DNA so that the overall HAC construct can function without the need to increase size via multimerization or acquisition of genomic DNA (Gambogi

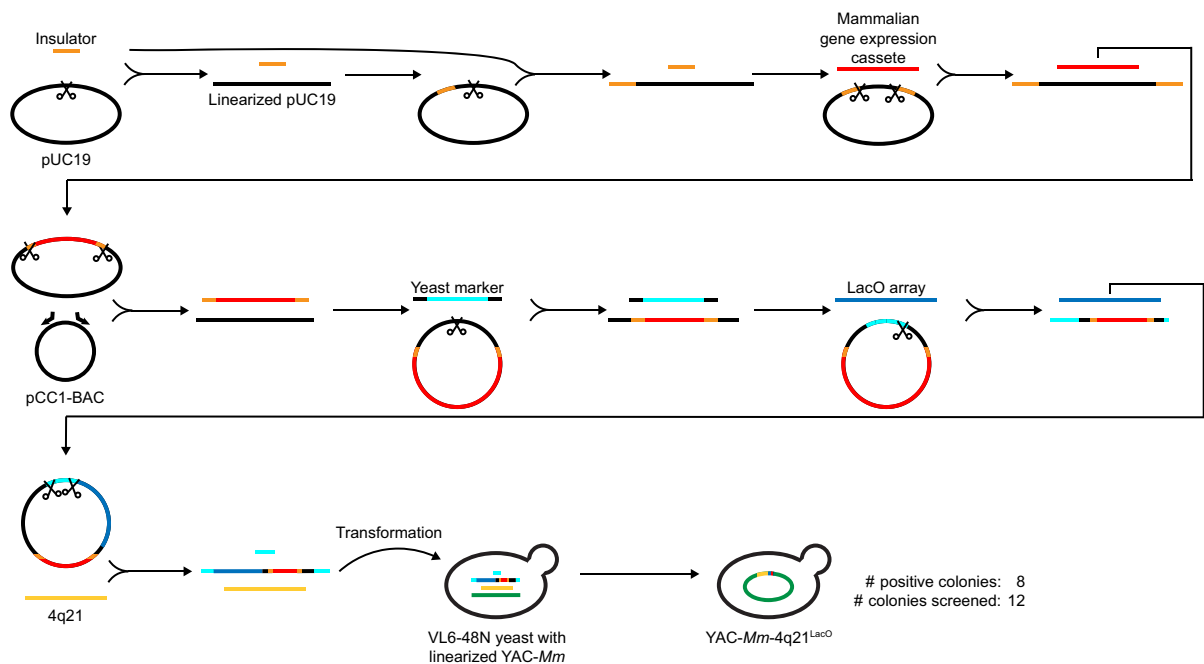


Fig. 1 Synthesizing a multi-module BAC for simplified and efficient TAR cloning to build a new YAC-Mm-4q21^{LacO}. Graphical depiction of cloning strategy to generate YAC-Mm-4q21^{LacO} in a single TAR step. First, a pair of tRNA dimers serving as insulators were inserted into the pUC19 vector. Subsequently, this vector was digested between the insulators and ligated to the *hEF1α* > *tdTomato*-P2A-*NeoR* expression cassette. Next, this recombinant plasmid was digested, and the insulator-flanked gene expression cassette was transferred to the linearized pCC1-BAC by isothermal assembly. This recombinant BAC was digested and mixed with a landing pad containing an auxotrophic marker, an 80-bp YAC hook with an internal restriction site, 40 bp of homology to the vector backbone housing the neocentromere-adjacent 4q21

locus, and 40 bp of homology to the vector backbone housing a 10 kb LacO array. Next, the vector was digested within the LacO landing pad and then isothermally assembled with the linearized LacO array, resulting in a BAC carrying both the LacO array and the gene expression cassette. In the final step, the BAC was digested in the YAC hook and in the 4q21 landing pad and co-transformed with the linearized 4q21 BAC and an sgRNA into YAC-Mm yeast constitutively expressing Cas9 to carry out TAR, generating the final YAC-Mm-4q21^{LacO} construct. Scissor symbols indicate restriction digestions at the corresponding cloning steps. Bent arrows indicate PCR primers sites for the corresponding cloning steps. Note that DNA fragments and vectors are not shown to scale for the sake of simplicity in diagramming out the overall strategy

et al. 2024; Logsdon et al. 2019). The HAC construct design requires the addition to the starting YAC of four functional modules: 1) a fluorescent protein to validate spheroplast fusion to human cells; 2) an antibiotic resistance gene to select for cells that harbor the HAC construct; 3) an array of bacterial Lac operator (LacO) sites for initial local assembly of CENP-A nucleosomes via a pulse of expression in recipient cells of a fusion protein of the CENP-A-specific histone chaperone, HJURP, and the Lac repressor (LacI); and 4) a 182 kb human chromosome segment (4q21) within which functional centromeric chromatin can form (Logsdon et al. 2019). A fifth module, a yeast auxotrophic selection marker, is also needed to select for yeast candidates that performed TAR.

Each of the intermediate steps in the assembly are built into standard plasmids (e.g., pUC19) or BACs (e.g., pCC1) (Fig. 1). The inserts include the synthesized mammalian expression cassette encoding both the fluorescent (tdTomato) and antibiotic (neomycin)-resistance proteins flanked by tRNA dimers (Ebersole et al. 2011) that are known to resist silencing mediated by pericentric heterochromatin (Blower and Karpen 2001; He and Brinkley 1996; Lee et al. 2013; Partridge et al. 2000), PCR amplicon (*C. albicans* URA3 gene), and digestion product from an existing vector (including the ~10 kb LacO array (Robinett et al. 1996)). At each step, homology “hooks” of ~40 bp were deployed for use with well-established and highly efficient isothermal assembly

cloning methodologies (Gibson et al. 2009). Indeed, all cloning steps through the final BAC prior to TAR cloning can be realistically accomplished by a single researcher in less than 3 weeks.

To expedite YAC assembly, this final BAC containing the engineered modules (with 4q21 hooks) and the BAC carrying 4q21 were linearized and co-transformed with an sgRNA into YAC-*Mm* yeast constitutively expressing Cas9 (Kannan et al. 2016) to carry out TAR, generating the final YAC-*Mm*-4q21^{LacO} construct. As expected with the relatively simple TAR reaction, hundreds of colonies survived auxotrophic selection. Further, 8/12 colonies screened by junctional PCR yielded products indicative of proper TAR assembly (Fig. 1). These results provided an early indication that our rapid HAC construct assembly strategy achieves the objective of efficiency and accuracy.

Genomic sequencing of the strain harboring the newly built YAC-*Mm*-4q21^{LacO}

We performed long-read Oxford nanopore technologies (ONT) whole-genome sequencing of a candidate yeast strain that had passed PCR-based diagnostic screening and acquired sufficient coverage to perform an unbiased whole-genome, de novo assembly. Alignment of reads back to the de novo-assembled YAC showed contiguous coverage (Fig. 2A). Alignment also reveals lower depth of coverage of the prokaryotic stuffer DNA (Fig. 2A). Given that the depth of coverage of other portions of the construct is comparable to that of the native yeast chromosomes (Fig. S1A), we concluded that lower depth of coverage in the stuffer DNA is due to an underrepresentation of many of the prokaryotic regions, possibly related to a technical bias against its extremely high A/T content. In support of this conclusion, an independently isolated TAR candidate exhibited a similar underrepresentation of much of the prokaryotic portions of the construct (Fig. S1B). The sequencing also showed that the second sequenced candidate is essentially identical to our initial candidate TAR-generated YAC, except for it having fewer LacO copies. Further analysis of the initial candidate YAC involved closer inspection of reads that span each junction, confirming seamless, on-target integration (Fig. 2B, C). We also leveraged the long-read sequencing modality to unambiguously evaluate LacO array

length and stability by selectively examining reads that spanned the entire array (i.e., had homology to unique sequence flanking the array). In doing so, we found a median LacO number of 259 among all reads with a large fraction (42/44) deviating by less than 10% (Fig. 2D), suggesting that the array is stable in culture despite the reported recombination proficiency of the yeast strain (Kouprina et al. 1998). Further, the genomic sequencing data were sufficient to assemble a circular 756 kb YAC that was mapped back to the designed reference sequence, confirming the absence of gross indels or rearrangements (Fig. 2E). This approach reveals a matrix associated with the human 4q21 region, a feature in our graph that we expected due to Alu elements that are highly abundant in 4q21 and throughout the human genome (Houck et al. 1979). It was further expected because this matrix matches that observed upon alignment of the 4q21-containing BAC (see Methods) used in the TAR cloning step to itself. The long-read sequencing data were also of sufficient quality to unambiguously assemble all native yeast chromosomes (Fig. 2F; Fig. S2). Together, our analysis of the strain harboring YAC-*Mm*-4q21^{LacO} revealed the desired TAR-mediated assembly alongside an intact, natural yeast genome.

HAC formation with YAC-*Mm*-4q21^{LacO} by spheroplast fusion and centromere seeding

We tested the ability of our newly built YAC-*Mm*-4q21^{LacO} construct to form HACs (Fig. 3) using the same general strategy (Fig. 3A) as reported with the prior version that had been generated through an independent cloning strategy (Gambogi et al. 2024). This involved YAC-*Mm*-4q21^{LacO} delivery to HT1080^{Dox-inducible mCherry-LacI-HJURP} or U2OS^{Dox-inducible mCherry-LacI-HJURP} cells via spheroplast fusion. An initial 48 h pulse of doxycycline (dox) was then applied to recipient cells before subjecting the cells to G418 selection. The polyclonal population of surviving cells was expanded to screen for HAC formation events by a combined fluorescence in situ hybridization (FISH) using probes targeting the *M. mycoides* stuffer that comprises ~550 kb of YAC-*Mm*-4q21^{LacO} and immunofluorescence (IF) to detect human CENP-A (Fig. 3A). As we anticipated based upon the earlier version of this YAC which forms functional HACs in cell populations after

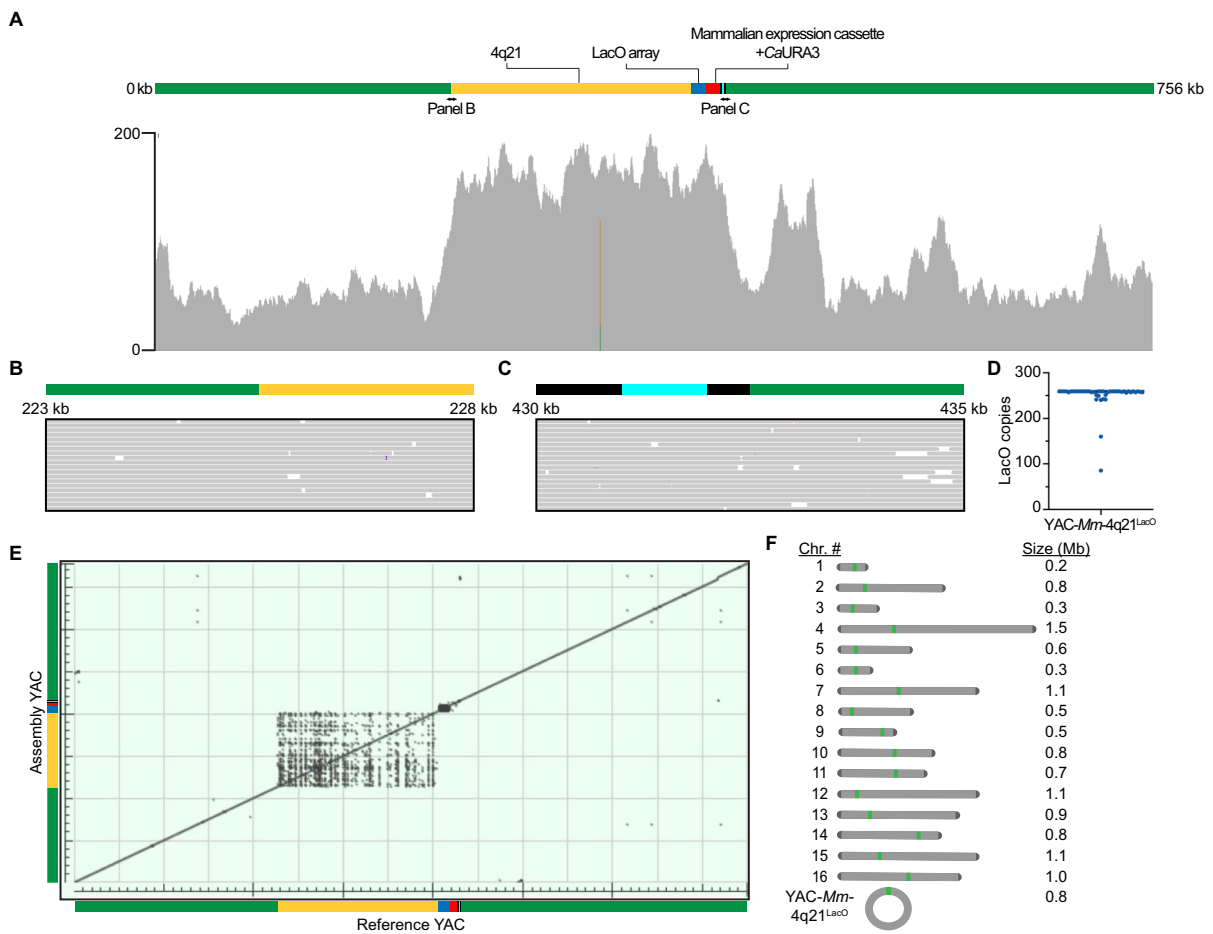


Fig. 2 Long-read DNA sequencing of the yeast genome harboring YAC-*Mm-4q21^{LacO}*. **A**) Integrated Genomics Viewer alignment track showing raw reads mapped to the de novo assembled YAC-*Mm-4q21^{LacO}*. Grey denotes a perfect match, copper denotes the unexpected presence of a guanine, and green denotes the unexpected presence of an adenine. In total, these unexpected instances impact 1 bp out of the entire 756 kb construct. **B**) Reads spanning a 5 kb region that includes the upstream junction of the BAC integration site. The purple bar represents a nucleotide insertion. Insertions that do not comprise consensus are a result of the error rate inherent to sequencing. **C**) Reads spanning a 5 kb region that includes the downstream junction of the BAC integration site. **D**) Distribution of LacO copies for each read spanning the entire

array. Bar denotes median LacO copy number of 259 for the 44 reads (individually represented by blue dots) analyzed. **E**) Dot plot generated by BLAST alignment of the de novo assembled YAC-*Mm-4q21^{LacO}* against the predicted recombinant YAC, referred to as “Reference YAC.” **F**) Graphical depiction of the YAC-*Mm-4q21^{LacO}* genome. All chromosomes except native chromosome 12 and lengths are derived from unbiased, genome assembly. De novo assembly failed to resolve the rDNA locus on chromosome 12, resulting in two contigs as previously reported (Salazar et al. 2017; Yue et al. 2017). Consequently, budding yeast chromosome 12 was manually assembled by joining the two contigs. Green denotes the yeast centromere, and dark grey denotes yeast telomeres

antibiotic selection but without the need to screen clones (Gambogi et al. 2024), HACs were observed and formation was stimulated by the early pulse of dox-inducible mCherry-LacI-HJURP expression in either HT1080 or U2OS backgrounds (Fig. 3B, C). In each chromosome spread containing a HAC, only a single HAC was detected. Further, we note that the

detected HACs (Fig. 3) were not of the large size associated with multimerization from smaller BACs, as in prior approaches with smaller BAC-based HAC formation approaches (Logsdon et al. 2019). Rather, they are small, consistent with our initial version of YAC-*Mm-4q21^{LacO}* that avoids rampant multimerization during HAC formation (Gambogi et al. 2024).

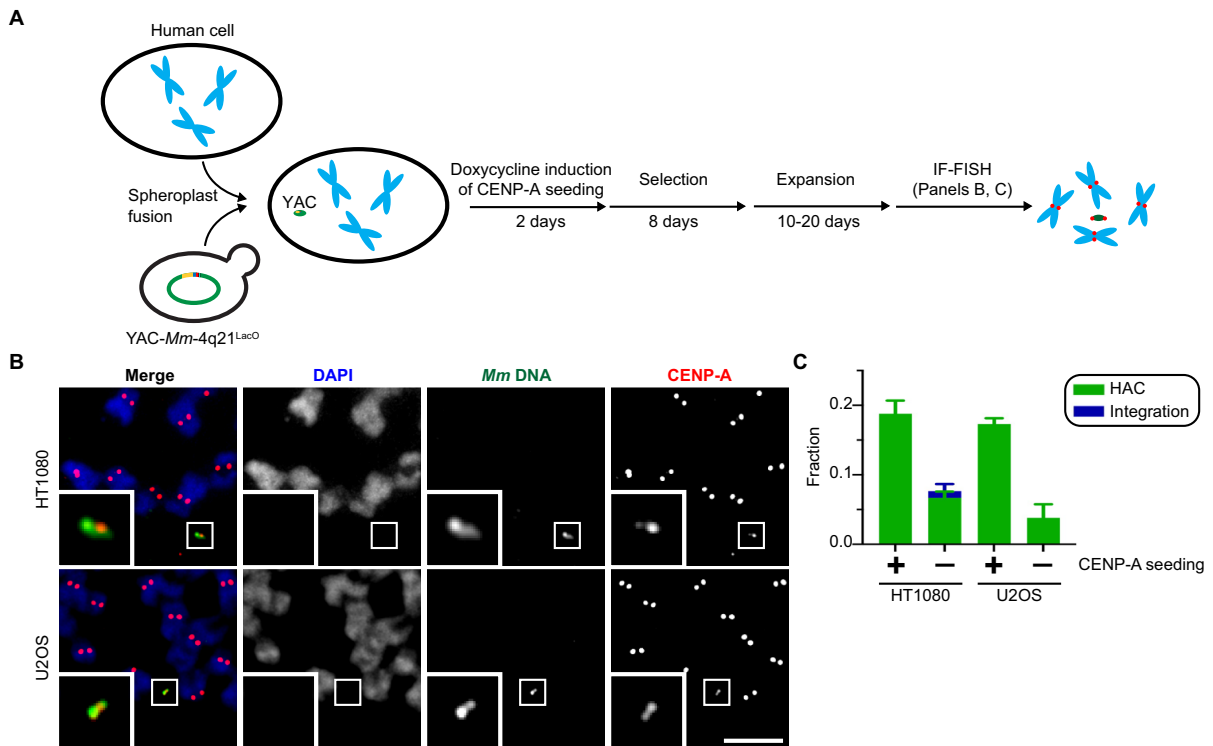


Fig. 3 YAC-Mm-4q21^{LacO} built via simplified TAR cloning forms functional HACs upon centromere seeding in human cells. **A**) Graphical representation of the HAC formation assay. **B**) Representative HACs formed after an initial pulse of mCherry-LacI-HJURP expression in HT1080 and U2OS cells. Inset, 5X magnification. Bar, 5 μ m. **C**) Quantification of HAC

formation frequency with and without doxycycline-mediated induction. The mean ($\pm$ standard error) is shown. The increase in HAC frequency was assessed using an unpaired, one-tailed T test. $P(\text{HT1080} + \text{dox} = \text{HT1080} - \text{dox}) < 0.001$, $n = 4$ (≥ 20 cells, replicate). $P(\text{U2OS} + \text{dox} = \text{U2OS} - \text{dox}) < 0.01$, $n = 3$ (≥ 20 cells, replicate)

While we measured HAC formation to assess function of the new assembly of YAC-Mm-4q21^{LacO}, we note that this study did not include assessment of its long-term stability in cell culture. Taken together, our HAC formation tests (Fig. 3) indicate efficient HAC formation with the newly constructed version of YAC-Mm-4q21^{LacO} is stimulated by epigenetic centromere seeding, in a similar manner as the original version (Gambogi et al. 2024).

Conclusions

HACs are attractive for their ability to test designed functional modules. YAC-based HAC construct assembly coupled with direct delivery to target human cells via spheroplast fusion promises to expand these modules in terms of their number and size due to their large carrying capacity. This promise raises new

challenges, including the ability to construct YACs of increasing complexity. In this study, we address one of these new challenges: the need for rapid and efficient HAC construct assembly. Without addressing this need, some TAR reactions are so inefficient that it is easy to imagine rounds of TAR attempts yielding no proper assemblies (i.e., no colonies surviving auxotrophic selection or no surviving colonies having correctly assembled recombinant YACs). What is needed is an approach that can be standardized to faithfully engineer HAC-precursor YACs. We achieved this with the new assembly of YAC-Mm-4q21^{LacO} by moving the majority of steps (Fig. 1) to in vitro assembly methodologies already proven to be rapid and efficient (Gibson et al. 2009). Multi-step cloning is accomplished in rapid in vitro assemblies in BACs. For new designs that one could envision, and based on our experience, one should be able to make global changes to BAC assemblies in less than a week.

Subsequent addition of LacO and 4q21 modules could also be added in less than a week, if that is preferred prior to TAR cloning. We note that the modules include potentially challenging DNA sequences, both in terms of size (which could hamper biochemical isolation and transformation into yeast, e.g., the 182 kb human chromosome 4q21 sequence) and repetitiveness (which could complicate the necessary recombination step in TAR cloning, e.g., the LacO array). The pipeline to couple YAC construction analysis to contemporary long-read sequencing methodologies is highly effective and is rendered rapid and efficient by well-established in-house and commercial microbial whole genome sequencing platforms. In sum, this advance is an important one among many that will enhance access to the design, synthesis, and testing of new HACs that address goals of diverse human chromosome engineering efforts.

Methods

BAC cloning

Four Integrated DNA Technologies (IDT) Ultramers (5'-TAGTGGTGAAGCATAGCTGCCTTCCAAGCAGTTGACCCGGGTTTCGATTCCCGGCCAACGCA GCGTGCCACGTTTTGCCTGGCGCCGCGTGCTAGCAAAGGTCCCTTTTACAGAGCAGTTTGAACACTGTTTTTGTAGGTTTTTCCAAGGGGATAATTATAGCTCATTGAGCTTTGGCGCGTGCTCCCTGGTGGTCTAG-3', 5'-CCAGCAGAGAGGAAGAAGAGAGGCTTCCCTGACCGGGAATCGAACCCGGGCGCGGCGGTGAGAGCGCCGAATCCTAACCACTAGACCACCAGGAGCACGCGC-3', 5'-ACGCGCCAATCCCATTGCAAAATTCTACAAAAGGAGTGTTTCCCAACTGCTCTATCAAGAGGAATGTTGCACACTGTGACCTGAATGCAAACATCACACGCGCCAGCAGAGAGGAAGAAGAGAGGCTTCCC-3', 5'-ACGCGTGATGGCTGCATTCCACACACAAGGTGGAACATTTCTCTTGATAGAGCAGTTTTGA AACACTCTTCCCGTAGAATCTGCTGGCGCGTGCTGCGCGTTGGTGGTATAGTGGTGAGCATAGCTGCCTTCCAAGC-3') were used in an overlap extension PCR reaction (primers: 5'-ACGCGCCAA TCCCAT-3' and 5'-ACGCGTGTGATGGCT-3') to generate a dimer of tRNA^{Gly} and tRNA^{Glu} genes (Ebersole et al. 2011). These dimers were serially

incorporated into a plasmid by isothermal assembly, resulting in a vector with an inverted dimer pair. Isothermal assembly was performed using NEB HiFi 2X master mix (NEB catalog #: E2621) according to the manufacturer's instructions. This vector and the plasmid carrying *hEF1α* > *tdTomato*-P2A-*NeoR*-SV40^{polyA} (synthesized by Twist Biosciences), were digested with enzymes SphI-HF and SalI-HF. The vector backbone including the insulators and the expression cassette insert were gel extracted using the QIAquick Gel Extraction Kit (Qiagen catalog #: 28704) and used in a two-fragment ligation with T4 DNA ligase (NEB catalog #: M0202) according to manufacturer instructions with the ligation being carried out at 16 °C overnight. The plasmid carrying the insulator-flanked gene expression cassette was validated by Plasmidsaurus sequencing.

The plasmid was digested by PvuII-HF to release the insulator-flanked expression cassette and isothermally assembled to AfeI-digested pCC1-BAC vector, which was PCR amplified at that site to contain 40 bp homology arms to the PvuII-released cassette. In a second round of DNA assembly, this new plasmid was digested with NotI and combined with the modular landing pad consisting of *C. albicans* *URA3* (PCR amplified from pAG60 plasmid (Addgene plasmid #: 35128) using primers 5'-ATAGGCCACTAGTGGATCTGATATCATCGATGAATTCGAG-3' and 5'-CTGTTTAGCTTGCCTCGTCC-3'), two 40-bp hooks for the microbial stuffer, two 40-bp hooks for 4q21 incorporation, and two 40-bp hooks for LacO array incorporation. This vector was then digested with AscI, and the plasmid carrying the LacO array (Logsdon et al. 2019) was digested with AseI and MfeI before combining these two fragments in a third round of isothermal assembly.

TAR cloning to generate YAC-*Mm*-4q21^{LacO} yeast

For TAR, the LacO-containing BAC was digested with PacI and NotI, and the BAC carrying the 4q21 locus (Logsdon et al. 2019) was digested with NotI. Both digests were phenol-chloroform extracted and mixed with an sgRNA generated by IDT (5'-mGmAmAGA UAAAAAUCCUCUAUCGUUUUAGAGCUAGAA AUAGCAAGUUAUAAAUAGGCUAGUCCGUUA UCAACUUGAAAAAGUGGCACCGAGUCG GUGCmUmUmUU-3') to target the Syn3B stuffer constituting YAC-*Mm*. Spheroplast transformation

was carried out as follows. YAC-*Mm* yeast (Bitten-court et al. 2024) (VL6-48N strain carrying YAC-*Mm* and constitutively expressing Cas9 (Kannan et al. 2016)) were grown in 50 ml synthetic minimal media -histidine to an OD₆₀₀ of 1.0. The culture was pelleted at 1,600×g for 3 min, washed with 50 ml sterile water, pelleted again, and resuspended in 20 ml 1 M sorbitol. The yeast were incubated on ice at 4 °C overnight. The next day, cells were harvested at 1,741×g for 3 min and resuspended in 20 ml SPE (1 M sorbitol, 7.89 mM Na₂HPO₄, 2.33 mM NaH₂PO₄, 10 mM EDTA pH 8.0). 40 µl of β-mercaptoethanol and 40 µl 10 mg/ml zymolyase-20 T were added to the suspension. The tube was inverted several times and then placed in a shaker to incubate at 30 °C, 50 rpm. After 20 min, the tube was inverted several times and again returned to the shaker to incubate for 20 min. 30 ml of 1 M sorbitol was added to the suspension before harvesting spheroplasts at 1,600×g for 5 min. Spheroplasts were washed with 50 ml 1 M sorbitol before harvesting at 1,600×g for 5 min. Spheroplasts were resuspended in 2.5 ml STC (1 M sorbitol, 10 mM Tris-Cl pH 7.5, 10 mM CaCl₂) and incubated at room temperature for 10 min. 200 µl of the suspension was added to an Eppendorf tube containing the DNA and sgRNA along with water to a total volume of 40 µl, and the suspension was incubated for 10 min at room temperature. 1 ml of PEG/CaCl₂ (20% w/v PEG8000, 10 mM CaCl₂, 10 mM Tris-Cl pH 7.5) was added and the solution mixed by inverting and flicking the Eppendorf tube. The solution was incubated at room temperature for 20 min. Transformed spheroplasts were harvested at 1,600×g for 8 min. The spheroplasts were suspended in 800 µl SOS medium (1 M sorbitol, 6.5 mM CaCl₂, 0.25% w/v bacto yeast extract, 0.5% w/v bacto peptone) and incubated at 30 °C for 1 h. The 800 µl solution was then pipetted onto 8 ml of 55 °C synthetic minimal top agar, inverted 3 times, and poured onto synthetic minimal sorbitol plates for recovery of transformants. -Ura prototrophs were patched onto synthetic minimal -histidine -uracil plates for genotyping.

Patches were screened by direct PCR without DNA extraction. Briefly, a QIAGEN Multiplex PCR Kit (cat # 206,145) was used to prepare a PCR master mix containing primers for either the left junction or the right junction. A small amount of cells from each of the patches were combined with 15 µl of the master mix and cycled per the instructions.

YAC-*Mm*-4q21^{LacO} genome assembly and analysis

YAC-*Mm*-4q21^{LacO} yeast were grown in 50 ml of synthetic minimal media (-histidine and -uracil) to OD₆₀₀=1.0. Yeast cells were pelleted at 500×g for 2 min. The pellet was resuspended in 0.5 ml of DNA/RNA shield (Zymo Research catalog #: R1100) before sending to Plasmidsaurus for ONT sequencing.

Raw reads from the ONT run were assembled using Flye (v. 2.9.2) in nano-hq mode while specifying the genome length as 13 Mb. Reads were mapped back to this de novo assembly using Minimap2 (v. 2.28-r1209), ensuring assembly contiguity. Samtools (v. 1.6) was used to filter out secondary and supplementary alignments, sort remaining primary aligned reads, and index aligned reads. The resulting alignment was visualized in the Integrated Genomics Viewer (v. 2.13.2). Sequencing and assembly statistics were reported using Samtools coverage (v 1.22) with the filtered alignment against the de novo assembly. Note that mapping quality takes the following into account: repetitive sequence composition of reference, base quality of mapping reads, the number of mismatches in aligned reads, and the number of candidate alignment hits on a reference. Statistics on all chromosomes but chromosome 12 are direct outputs of Samtools coverage. For chromosome 12, which the de novo assembly failed to construct as a single contig, a weighted average of the values reported for the two constituent contigs, excluding the rDNA locus, is shown in Fig. S2.

To evaluate LacO array length, Seqtk (v. 1.4-r122) was used to retrieve pass-filter reads mapping to 300 unique nucleotides upstream and downstream of the LacO array. The length of each LacO array, considered the distance between the first bp of the first monomer and the last bp of the last monomer, was divided by the number of bp per monomer (39.85, including the monomer and variable-length intervening sequence) to generate the number of LacO copies for each read. The assembled sequence of YAC-*Mm*-4q21^{LacO} is provided as a supplemental data file.

Spheroplast fusion

YAC-*Mm*-4q21^{LacO} yeast were grown in synthetic minimal media (-histidine -uracil) to OD₆₀₀=0.8–1.0 before harvesting at 1,741×g for 3 min. Yeast were resuspended in 20 ml 1 M sorbitol and incubated

overnight at 4 °C. Yeast were harvested at $1,741\times g$ for 3 min and resuspended in 20 ml SPE medium. 40 μ l β -mercaptoethanol and 100 μ l 10 mg/ml zymolyase-20 T were added to the yeast solution before gently inverting and then incubating at 37 °C in a shaker set to 25 rpm for 1 h. 30 ml ice cold 1 M sorbitol was then added to the spheroplasts before pelleting at $627\times g$, 18 min, 4 °C. The spheroplasts were washed in 50 ml ice cold 1 M sorbitol before harvesting at $627\times g$, 25 min, 4 °C. The pellet was resuspended in 1 ml STC, and the solution was allowed to incubate at room temperature for 10–30 min before addition to the HT1080^{Dox-inducible mCherry-LacI-HJURP} or U2OS^{Dox-inducible mCherry-LacI-HJURP} cells. The cells were then processed along with the spheroplasts. ~4 h before yeast were pelleted, HT1080^{Dox-inducible mCherry-LacI-HJURP} cells or U2OS^{Dox-inducible mCherry-LacI-HJURP} cells (~60–80% confluent) were supplied with fresh DMEM containing 10% certified tetracycline-free fetal bovine serum (Takara catalog #: 631,106), 100 U/ml penicillin, and 100 μ g/ml streptomycin (hereafter referred to as complete DMEM) and 50 μ M S-trityl-L-cysteine (STLC) (Fisher catalog #: AAL1438406). After 5–6 h incubation in 50 μ M STLC, cells were washed with PBS, trypsinized, and counted using a hemacytometer. Cells were pelleted and resuspended in PBS to a concentration of 6×10^5 cells/ml. 3×10^5 HT1080^{Dox-inducible mCherry-LacI-HJURP} cells or U2OS^{Dox-inducible mCherry-LacI-HJURP} cells were mixed with 9×10^7 YAC-Mm-4q21^{LacO} spheroplasts and incubated at room temperature for 5 min. The mixture was pelleted at $1,700\times g$ for 30 s in a microcentrifuge. The supernatant was aspirated and the pellet resuspended in a 45% PEG2000, 10% DMSO solution in 75 mM HEPES pH 7.5. The suspension was incubated at room temperature for 5 min. The fusion was quenched by adding 1 ml serum/antibiotic-free DMEM. Fused cells were pelleted at $1,700\times g$ for 30 s in a microcentrifuge. Cells were resuspended in 1 ml complete DMEM and 500 μ l aliquoted into a 6-well plate containing 2 ml complete DMEM for the HAC formation assay and 500 μ l seeded onto a coverslip for analysis of fusion efficiency at 72 h. After 4 h (when cells have adhered), media is swapped with 3 ml complete DMEM $\pm$ 2 μ g/ml doxycycline. After 16 h, cells are again supplied with 3 ml complete DMEM $\pm$ 2 μ g/ml doxycycline. Cells are allowed to incubate for 32 h before washing with PBS, trypsinizing, and transferring to a 10-cm plate with complete DMEM + 325 μ g/

ml G418 in the case of HT1080^{Dox-inducible mCherry-LacI-HJURP} cells or 267 μ g/ml G418 in the case of U2OS^{Dox-inducible mCherry-LacI-HJURP} cells. Selection is maintained for 8 days before reducing the G418 concentration to the 150 μ g/ml maintenance concentration, in the case of HT1080^{Dox-inducible mCherry-LacI-HJURP} cells, or 100 μ g/ml, in the case of U2OS^{Dox-inducible mCherry-LacI-HJURP} cells, for subsequent expansion. Note that the media is refreshed in the coverslip wells alongside the HAC formation assay wells but in the absence of doxycycline to enable detection of fluorescent signal derived from spheroplast fusion. Evidence of fusion events is discerned by red fluorescent, yeast-derived cytoplasmic foci.

IF-FISH on metaphase spreads

70–80% confluent human cells were fed with complete DMEM containing 50 μ M STLC and incubated at 37 °C for 2–4 h. Mitotic cells were blown off using a transfer pipet and pelleted at $415\times g$ for 5 min. Cells were resuspended in 500 μ l PBS before counting with a hemacytometer. The cells were again pelleted at $415\times g$ for 5 min before dropwise resuspending the cells in hypotonic buffer consisting of a 1:1:1 ratio of 75 mM KCl, 0.8% $\text{Na}_3\text{C}_6\text{H}_5\text{O}_7$, and a solution consisting of 1.5 mM MgCl_2 and 3 mM CaCl_2 to a concentration of 5×10^5 cells/ml. Cells were allowed to swell at room temperature for 15 min before diluting to 5×10^4 cells/ml in 500 μ l total hypotonic buffer. Cells were cytospun at 1500 rpm, acceleration 9, for 5 min in a Shandon Cytospin 4 centrifuge. Cells were allowed to dry for 1 min before permeabilizing with KCM buffer (10 mM Tris-Cl pH 7.7, 120 mM KCl, 20 mM NaCl, and 0.1% triton X-100) for 15 min. Spreads were blocked in IF block (2% FBS, 2% BSA, 0.1% tween-20, and 0.02% NaN_3) for 20 min before incubating with α CENP-A mAb (ENZO catalog #: ADI-KAM-CC006) diluted to 1 μ g/ml in IF block for 45 min at room temperature. Spreads were washed 3 times in KCM for 5 min each before incubating in Cy3-conjugated donkey anti-mouse IgG (Jackson ImmunoResearch catalog #: 712-165-151) diluted to 3.5 μ g/ml in IF block for 25 min. Spreads were again washed 3 times in KCM for 5 min before fixing in 4% paraformaldehyde for 10 min at room temperature. Spreads were washed 3 times with dH_2O before proceeding with FISH.

FISH probe was prepared by nick translation (Roche catalog #: 10976776001) using biotin-16-dUTP (Millipore catalog #: 11093070910) and *M. mycoides* gDNA as template. Unincorporated nucleotides were removed using Amersham Micro-Spin G-50 columns (Cytiva catalog #: 27533002). 300 ng of probe/slide was ethanol precipitated along with 20 µg of salmon sperm DNA (Invitrogen catalog #: 15632011) and 10 µg of human Cot1 DNA (Invitrogen catalog #: 15279011). Precipitated DNA was resuspended in 50% formamide, 10% dextran sulfate in 2X saline-sodium citrate (SSC; Corning catalog #: 46-020-CM) and denatured at 78 °C for 10 min before being placed in 37 °C for at least 20 min for pre-hybridization. 300 ng of probe was added to the RNase-treated spreads, and hybridization was performed at 37 °C for 16 h. Slides were washed 2 times in 50% formamide, 2X SSC at 37 °C for 5 min before washing slides 2 times in 2X SSC at 37 °C for 5 min. Slides were blocked with 2.5% milk in 4X SSC+0.1% tween-20 for 10 min and then incubated in NeutrAvidin-FITC (ThermoFisher catalog #: 31006) diluted to 25 µg/ml with 2.5% milk in 4X SSC+0.1% tween-20 at 37 °C for 1 h. Spreads were washed 3 times in 4X SSC+0.1% tween-20 at 45 °C at 45 °C for 2 min, followed by one wash in PBS. Spreads were then incubated in PBS supplemented with 1 µg/ml DAPI for 10 min at room temperature, washed again in PBS, and then water before mounting with Vectashield. Slides were imaged on an inverted epifluorescence microscope (Leica DMI6000B) with a 100×oil-immersion objective and the Leica K8 Scientific CMOS camera. Images were blind deconvolved within LAS X (v. 3.6.0.20104) while resizing from 12-bit depth to 16-bit depth. Channel contrast and image cropping was carried out in ImageJ (v. 1.52 k) before assembling figures in Adobe Illustrator (v. 29.7.1). The “HAC” designation was given to objects with FISH signal < 2 µm in diameter and two co-localized CENP-A foci. The “integration” designation was given to events with FISH signal located at the same site on each sister chromatid of a native chromosome and distal from the centromere.

Acknowledgements We thank our UPenn colleagues G.A. Logsdon and S-C. Chuang for advice related to ONT sequencing and analysis workflows and G.A. Logsdon for comments on the manuscript. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health (NIH). This manuscript is the

result of funding in whole or in part by the NIH. It is subject to the NIH Public Access Policy. Through acceptance of this federal funding, NIH has been given a right to make this manuscript publicly available in PubMed Central upon the Official Date of Publication, as defined by NIH.

Author contributions G.J.B., D.G.G., and B.E.B. conceived the study. G.J.B., D.G.G., P.K.A., and P.K. performed the experiments. G.J.B., P.K.A., P.K., and B.E.B. analyzed data. G.J.B. and B.E.B. wrote the paper. All authors edited the paper. J.I.G. and B.E.B. directed the research.

Funding Supported by NIH grants HG012445 (J.I.G. and B.E.B.) and CA265794 (B.E.B.). G.J.B. was supported by the Penn Cell and Molecular Biology Training Program: GM007229.

Data availability Sequencing data are available on the Sequence Read Archive (accession #: PRJNA1322447). Remaining data required to evaluate the conclusions of this report may be found in this paper. Reagents used and described are available from the authors upon request.

Declarations

Competing interests G.J.B., J.I.G., and B.E.B. are inventors on a patent application submitted by the University of Pennsylvania related to this work.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Amano M, Suzuki A, Hori T, Backer C, Okawa K, Cheeseman IM, Fukagawa T (2009) The CENP-S complex is essential for the stable assembly of outer kinetochore structure. *J Cell Biol* 186(2):173–182. <https://doi.org/10.1083/jcb.200903100>
- Barnhart MC, Kuich PHJL, Stellfox ME, Ward JA, Bassett EA, Black BE, Foltz DR (2011) HJURP is a CENP-A chromatin assembly factor sufficient to form a functional de novo kinetochore. *J Cell Biol* 194(2):229–243. <https://doi.org/10.1083/jcb.201012017>

- Blower MD, Karpen GH (2001) The role of *Drosophila* CID in kinetochore formation, cell-cycle progression and heterochromatin interactions. *Nat Cell Biol* 3(8):730–739. <https://doi.org/10.1038/35087045>
- Chen C-C, Dechassa ML, Bettini E, Ledoux MB, Belisario C, Heun P, Luger K, Mellone BG (2014) CAL1 is the *Drosophila* CENP-A assembly factor. *J Cell Biol* 204(3):313–329. <https://doi.org/10.1083/jcb.201305036>
- Bittencourt DM de C, Brown DM, Assad-Garcia N, Romero MR, Sun L, Palhares de Melo LAM, Freire M, Glass JI (2024) Minimal bacterial cell JCVI-syn3B as a chassis to investigate interactions between bacteria and mammalian cells. *ACS Synth Biol* 13(4):1128–1141. <https://doi.org/10.1021/acssynbio.3c00513>
- Earnshaw WC, Rothfield N (1985) Identification of a family of human centromere proteins using autoimmune sera from patients with scleroderma. *Chromosoma* 91(3–4):313–321. <https://doi.org/10.1007/bf00328227>
- Ebersole TA, Ross A, Clark E, McGill N, Schindelhauer D, Cooke H, Grimes B (2000) Mammalian artificial chromosome formation from circular aliphoid input DNA does not require telomere repeats. *Hum Mol Genet* 9(11):1623–1631. <https://doi.org/10.1093/hmg/9.11.1623>
- Ebersole T, Kim J-H, Samoshkin A, Kouprina N, Pavlicek A, White RJ, Larionov V (2011) tRNA genes protect a reporter gene from epigenetic silencing in mouse cells. *Cell Cycle* 10(16):2779–2791. <https://doi.org/10.4161/cc.10.16.17092>
- Foltz DR, Jansen LET, Black BE, Bailey AO, Yates JR, Cleveland DW (2006) The human CENP-A centromeric nucleosome-associated complex. *Nat Cell Biol* 8(5):458–469. <https://doi.org/10.1038/ncb1397>
- Gambogi CW, Birchak GJ, Mer E, Brown DM, Yankson G, Kixmoeller K, Gavade JN, Espinoza JL, Kashyap P, Dupont CL, Logsdon GA, Heun P, Glass JI, Black BE (2024) Efficient formation of single-copy human artificial chromosomes. *Science* 383(6689):1344–1349. <https://doi.org/10.1126/science.adj3566>
- Gibson DG, Benders GA, Andrews-Pfannkuch C, Denisova EA, Baden-Tillson H, Zaveri J, Stockwell TB, Brownley A, Thomas DW, Algire MA, Merryman C, Young L, Noskov VN, Glass JI, Venter JC, Hutchison CA, Smith HO (2008a) Complete chemical synthesis, assembly, and cloning of a *Mycoplasma genitalium* genome. *Science* 319(5867):1215–1220. <https://doi.org/10.1126/science.1151721>
- Gibson DG, Benders GA, Axelrod KC, Zaveri J, Algire MA, Moodie M, Montague MG, Venter JC, Smith HO, Hutchison CA (2008b) One-step assembly in yeast of 25 overlapping DNA fragments to form a complete synthetic *Mycoplasma genitalium* genome. *Proc Natl Acad Sci U S A* 105(51):20404–20409. <https://doi.org/10.1073/pnas.0811011106>
- Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 6(5):343–345. <https://doi.org/10.1038/nmeth.1318>
- Harrington JJ, Bokkelen GV, Mays RW, Gustashaw K, Willard HF (1997) Formation of *de novo* centromeres and construction of first-generation human artificial microchromosomes. *Nat Genet* 15(4):345–355. <https://doi.org/10.1038/ng0497-345>
- Hayden KE, Strome ED, Merrett SL, Lee H-R, Rudd MK, Willard HF (2013) Sequences associated with centromere competency in the human genome. *Mol Cell Biol* 33(4):763–772. <https://doi.org/10.1128/MCB.01198-12>
- He D, Brinkley BR (1996) Structure and dynamic organization of centromeres/prekinetochores in the nucleus of mammalian cells. *J Cell Sci* 109(11):2693–2704. <https://doi.org/10.1242/jcs.109.11.2693>
- Hori T, Amano M, Suzuki A, Backer CB, Welburn JP, Dong Y, McEwen BF, Shang W-H, Suzuki E, Okawa K, Cheeseman IM, Fukagawa T (2008) CCAN makes multiple contacts with centromeric DNA to provide distinct pathways to the outer kinetochore. *Cell* 135(6):1039–1052. <https://doi.org/10.1016/j.cell.2008.10.019>
- Hori T, Shang W-H, Takeuchi K, Fukagawa T (2012) The CCAN recruits CENP-A to the centromere and forms the structural core for kinetochore assembly. *J Cell Biol* 200(1):45–60. <https://doi.org/10.1083/jcb.201210106>
- Houck CM, Rinehart FP, Schmid CW (1979) A ubiquitous family of repeated DNA sequences in the human genome. *J Mol Biol* 132(3):289–306. [https://doi.org/10.1016/0022-2836\(79\)90261-4](https://doi.org/10.1016/0022-2836(79)90261-4)
- Ikeno M, Grimes B, Okazaki T, Nakano M, Saitoh K, Hoshino H, McGill NI, Cooke H, Masumoto H (1998) Construction of YAC-based mammalian artificial chromosomes. *Nat Biotechnol* 16(5):431–439. <https://doi.org/10.1038/nbt0598-431>
- Kannan K, Tsvetanova B, Chuang R-Y, Noskov VN, Assad-Garcia N, Ma L, Hutchison CA III, Smith HO, Glass JI, Merryman C, Venter JC, Gibson DG (2016) One step engineering of the small-subunit ribosomal RNA using CRISPR/Cas9. *Sci Rep* 6(1):30714. <https://doi.org/10.1038/srep30714>
- Kixmoeller K, Allu PK, Black BE (2020) The centromere comes into focus: from CENP-A nucleosomes to kinetochore connections with the spindle. *Open Biol*. <https://doi.org/10.1098/rsob.200051>
- Kouprina N, Annab L, Graves J, Afshari C, Barrett JC, Resnick MA, 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 U S A* 95(8):4469–4474. <https://doi.org/10.1073/pnas.95.8.4469>
- Kouprina N, Samoshkin A, Erliandri I, Nakano M, Lee H-S, Fu H, Iida Y, Aladjem M, Oshimura M, Masumoto H, Earnshaw WC, Larionov V (2012) Organization of synthetic aliphoid DNA array in human artificial chromosome (HAC) with a conditional centromere. *ACS Synth Biol* 1(12):590–601. <https://doi.org/10.1021/sb3000436>
- Lee NC, Kononenko AV, Lee H-S, Tolkunova EN, Liskovych MA, Masumoto H, Earnshaw WC, Tomilin AN, Larionov V, Kouprina N (2013) Protecting a transgene expression from the HAC-based vector by different chromatin insulators. *Cell Mol Life Sci* 70(19):3723–3737. <https://doi.org/10.1007/s00018-013-1362-9>
- Logsdon GA, Gambogi CW, Liskovych MA, Barrey EJ, Larionov V, Miga KH, Heun P, Black BE (2019) Human artificial chromosomes that bypass centromeric DNA. *Cell* 178(3):624–639. <https://doi.org/10.1016/j.cell.2019.06.006>

- Maloney KA, Sullivan LL, Matheny JE, Strome ED, Merrett SL, Ferris A, Sullivan BA (2012) Functional epialleles at an endogenous human centromere. *Proc Natl Acad Sci U S A* 109(34):13704–13709. <https://doi.org/10.1073/pnas.1203126109>
- Mendiburo MJ, Padeken J, Fülöp S, Schepers A, Heun P (2011) *Drosophila* CENH3 is sufficient for centromere formation. *Science* 334(6056):686–690. <https://doi.org/10.1126/science.1206880>
- Musacchio A, Desai A (2017) A molecular view of kinetochore assembly and function. *Biol* 6(1):5. <https://doi.org/10.3390/biology6010005>
- Ohzeki J, Nakano M, Okada T, Masumoto H (2002) CENP-B box is required for de novo centromere chromatin assembly on human alphoid DNA. *J Cell Biol* 159(5):765–775. <https://doi.org/10.1083/jcb.200207112>
- Okada M, Cheeseman IM, Hori T, Okawa K, McLeod IX, Yates JR Iii, Desai A, Fukagawa T (2006) The CENP-H–I complex is required for the efficient incorporation of newly synthesized CENP-A into centromeres. *Nat Cell Biol* 8(5):446–457. <https://doi.org/10.1038/ncb1396>
- Okada T, Ohzeki J, Nakano M, Yoda K, Brinkley WR, Larionov V, Masumoto H (2007) CENP-B controls centromere formation depending on the chromatin context. *Cell* 131(7):1287–1300. <https://doi.org/10.1016/j.cell.2007.10.045>
- Palmer DK, O'Day K, Wener MH, Andrews BS, Margolis RL (1987) A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones. *J Cell Biol* 104(4):805–815. <https://doi.org/10.1083/jcb.104.4.805>
- Partridge JF, Borgström B, Allshire RC (2000) Distinct protein interaction domains and protein spreading in a complex centromere. *Genes Dev* 14(7):783–791
- Robinett CC, Straight A, Li G, Wilhelm C, Sudlow G, Murray A, Belmont AS (1996) In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. *J Cell Biol* 135(6):1685–1700. <https://doi.org/10.1083/jcb.135.6.1685>
- Rudd MK, Mays RW, Schwartz S, Willard HF (2003) Human artificial chromosomes with alpha satellite-based de novo centromeres show increased frequency of nondisjunction and anaphase lag. *Mol Cell Biol* 23(21):7689–7697. <https://doi.org/10.1128/MCB.23.21.7689-7697.2003>
- Salazar AN, Gorter de Vries AR, van den Broek M, Wijsman M, de la Torre Cortés P, Brickwedde A, Brouwers N, Daran J-MG, Abeel T (2017) Nanopore sequencing enables near-complete de novo assembly of *Saccharomyces cerevisiae* reference strain CEN.PK113–7D. *FEMS Yeast Res* 17(7). <https://doi.org/10.1093/femsyr/fox074>
- Schueler MG, Higgins AW, Rudd MK, Gustashaw K, Willard HF (2001) Genomic and genetic definition of a functional human centromere. *Science* 294(5540), 109–115.
- Yue J-X, Li J, Aigrain L, Hallin J, Persson K, Oliver K, Bergström A, Coupland P, Warringer J, Lagomarsino MC, Fischer G, Durbin R, Liti G (2017) Contrasting evolutionary genome dynamics between domesticated and wild yeasts. *Nat Genet* 49(6):913–924. <https://doi.org/10.1038/ng.3847>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.