b. In order of importance, list of ten best papers of the candidate, highlighting the important discoveries/contributions

1. Yadav V, Sun S, Billmyre RB, Thimmappa BC, Shea T, Linter R, Bakkeren G, Cuomo CA, Heitman J, **Sanyal K* (2018)** RNAi is a critical determinant of centromere evolution in closely related fungi.

Proc Natl Acad Sci USA 115: 3108 – 3113 (*F1000 recommended*)

Identification of centromeres in three closely related species of the *Cryptococcus* species complex reveals that RNAi is a major determinant of the centromere length maintenance. The RNAi-proficient *Cryptococcus* species maintain full-length retrotransposons at the centromeres which are longer than those of the RNAi-deficient species that lost all full-length retrotransposons at the centromere. This has been further evidenced by an experimental evolution experiment using an RNAi-mutant of the RNAi-proficient species. The nominee was involved in the identification of centromeres in all three species of *Cryptococcus* by ChIP-sequencing, analysis of transposons and experimental evolution experiments.

2. Padmanabhan S, Thakur J, Siddharthan R, **Sanyal K*. (2008)** Rapid evolution of Cse4p-rich centromeric DNA sequences in closely related pathogenic yeasts, *Candida albicans* and *Candida dubliniensis*.

Proc Natl Acad Sci USA 105:19797-19802 (*F1000 recommended*)

In this study, the nominee's group identified the centromeres of *Candida dubliniensis*, a closely related species of *Candida albicans*. The centromere properties are shown to be similar to those of *Candida albicans*. The striking finding was the rapid evolution of the centromere DNA sequence as it was through this study the authors established that centromeres are the most rapidly evolving regions in the genome of these two very closely related species.

3. Sridhar S, Hori T, Nakagawa R, Fukagawa T*, **Sanyal K*** (2021) Bridgin connects the outer kinetochore to centromeric chromatin.

Nature Communications 12: 146. doi: 10.1038/s41467-020-20161-9 (*F1000 recommended*)

The microtubule-binding outer kinetochore is coupled to centromeric chromatin through CENP-C^{Mif2}, CENP-T^{Cnn1}, and CENP-U^{Ame1} linker pathways originating from the constitutive centromere associated network (CCAN) of the inner kinetochore. This study reported a recurrent loss of most CCAN components, including the above-mentioned linkers among several species of Basidiomycota. The kinetochore interactome was studied using the human fungal pathogen *Cryptococcus neoformans* as a model, where a forkhead-associated domain-containing protein "bridgin" was identified as a kinetochore component along with other predicted kinetochore proteins. In vivo and in vitro functional analyses of bridgin reveal its ability to connect the outer kinetochore with centromeric chromatin to ensure accurate chromosome segregation. Unlike established CCAN-based linkers, bridgin is recruited at the outer kinetochore establishing its role as a distinct family of kinetochore proteins. The presence of bridgin homologs in non-fungal lineages suggests an ancient divergent strategy exists to bridge the outer kinetochore with centromeric chromatin.

4. Thakur J, **Sanyal K*. (2013)** Efficient neocentromere formation is suppressed by gene conversion to maintain centromere function at native physical chromosomal loci in *Candida albicans*.

Genome Research 23:638-652.

This landmark study identifies neocentromeres in *Candida albicans* and reveals that the genomic location rather than the DNA sequence determines the centromere location. This study also demonstrates that the wild-type centromere locus could be acquired by the neocentric chromosome via gene

conversion, and as a result, the neocentromere mark is lost. All the experiments and analysis of data were performed in the nominee's laboratory.

5. Sankaranarayanan SR, Ianiri G, Coelho MA, Reza MH, Thimmappa BC, Ganguly P, Vadnala RN, Sun S, Siddharthan R, Tellgren-Roth C, Dawson TL Jnr, Heitman J*, **Sanyal K*** (2020) Loss of centromere function drives karyotype evolution in closely related *Malassezia* species.

eLife 9: e53944. doi: 10.7554/eLife.53944 (*highlighted in an eLife digest*).

Millions of yeast, bacteria and other microbes live in or on the human body. A type of yeast known as *Malassezia* is one of the most abundant microbes living on our skin. Generally, *Malassezia* does not cause symptoms in humans but are associated with dandruff, dermatitis and other skin conditions in susceptible individuals. There are 18 closely related species of *Malassezia* and all have an unusually small sized genome as compared to other yeasts. The number of chromosomes in different *Malassezia* species varies between six and nine.

A region of each chromosome known as the centromere is responsible for ensuring that equal numbers of chromosomes are passed on to their offspring. This means that any defects in centromeres can lead to the daughter yeast cells inheriting unequal numbers of chromosomes. Changes in chromosome number can drive the evolution of new species, but it remains unclear if and how centromere loss may have contributed to the evolution of *Malassezia* species. Biochemical, molecular genetic, and comparative genomic approaches were used to study the chromosomes of *Malassezia* species. The experiments revealed that nine *Malassezia* species had centromeres that shared common features such as being AT-rich. We propose that AT-rich nature make the centromeres more fragile leading to occasional breaks. This may have contributed to the loss of centromeres in some *Malassezia* cells and helped new species to evolve with fewer chromosomes. A better understanding of how *Malassezia* organize their genetic material should enable in-depth studies of how these yeasts interact with their human hosts and how they contribute to skin disease, cancer, Crohn's disease and other health conditions. More broadly, these findings may help scientists to better understand how changes in chromosomes cause new species to evolve.

.

6. Guin K, Chen Y, Mishra R, Muzaki SRBM, Thimmappa BC, O'Brien C, Butler G, Sanyal A*, **Sanyal K*** (2020) Spatial proximity of homologous centromere DNA sequences facilitated karyotype diversity andseeding of evolutionary new centromeres.

eLife 9: e58556 (*F1000 recommended*).

Centromeres of *Candida albicans* form on unique and different DNA sequences but a closely related species, *Candida tropicalis*, possesses homogenized inverted repeat (HIR)-associated centromeres. To investigate the mechanism of centromere type transition, we improved the fragmented genome assembly and constructed a chromosome-level genome assembly of *C. tropicalis* by employing PacBio sequencing, chromosome conformation capture sequencing (3C-seq), chromoblot, and genetic analysis of engineered aneuploid strains. Further, we analyzed the 3D genome organization using 3C-seq data, which revealed spatial proximity among the centromeres as well as telomeres of seven chromosomes in *C. tropicalis*. Intriguingly, we observed evidence of inter-centromeric translocations in the common ancestor of *C. albicans* and *C. tropicalis*. Identification of putative centromeres in closely related *Candida sojae*, *Candida viswanathii* and *Candida parapsilosis* indicates loss of ancestral HIR-associated centromeres and establishment of evolutionary new centromeres (ENCs) in *C. albicans*. We propose that spatial proximity of the homologous centromere DNA sequences facilitated karyotype rearrangements and centromere type transitions in human pathogenic yeasts of the CUG-Ser1 clade.

7. Sreekumar L, Kumari K, Guin K, Bakshi A, Varshney N, Thimmappa BC, Narlikar L, Padinhateeri R,Siddharthan R, **Sanyal K*** (2021) Orc4 spatiotemporally stabilizes centromeric chromatin.

Genome Research 31:607- 621. doi:10.1101/gr.265900.120.

DNA replication origins, origin binding proteins, and replication timing of centromere DNA are important determinants of centromere function. The epigenetically regulated regional centromeres in the budding yeast *Candida albicans* have unique DNA sequences that replicate earliest in every chromosome and are clustered throughout the cell cycle. In this study, the genome-wide occupancy of the replication initiation protein Orc4 reveals its abundance at all centromeres in *C. albicans*. Orc4 is associated with four different DNA sequence motifs, one of which coincides with tRNA genes (tDNA) that replicate early and cluster together in space. We simulate a polymer model of chromosomes of *C. albicans* and propose that the early replicating and highly enriched Orc4-bound sites preferentially localize around the clustered kinetochores. We also observe that Orc4 is constitutively localized to centromeres, and both Orc4 and the helicase Mcm2 are essential for cell viability and CENPA stability in *C. albicans*. Finally, we show that new molecules of CENPA are recruited to centromeres during late anaphase/telophase, which coincides with the stage at which the CENPA-specific chaperone Scm3 localizes to the kinetochore. We propose that the spatiotemporal localization of Orc4 within the nucleus, in collaboration with Mcm2 and Scm3, maintains centromeric chromatin stability and CENPA recruitment in *C. albicans*.

8. Narayanan A, Vadnala RN, Ganguly P, Selvakumar P, Rudramurthy SM, Prasad R, Chakrabarti A, Siddharthan R, **Sanyal K*** (2021) Functional and comparative analysis of centromeres reveals clade-specific rearrangements in *Candida auris* and a chromosome number change in related species.

mBio 12: e00905-21 (*two accompanied commentaries published in mBio)*

Candida auris, the killer fungus, emerged as different geographical clades, exhibiting multidrug resistance and high karyotype plasticity. Chromosomal rearrangements are known to play key roles in the emergence of new species, virulence, and drug resistance in pathogenic fungi. Centromeres, the genomic loci where microtubules attach to separate the sister chromatids during cell division, are known to be hot spots of breaks and downstream rearrangements. We identified the centromeres in *C. auris* and related species to study their involvement in the evolution and karyotype diversity reported in *C. auris* We report conserved centromere features in 10 related species and trace the events that occurred at the centromeres during evolution. We reveal a centromere inactivation-mediated chromosome number change in these closely related species. We also observe that one of the geographical clades, the East Asian clade, evolved along a unique trajectory, compared to the other clades and related species.

9. Rai LS, Singha R, Sanchez H, Chakraborty T, Chand B, Bachellier-Bassi S, Chowdhury S, d'Enfert C, Andes DR, **Sanyal K*** (2019) The *Candida albicans* biofilm gene circuit modulated at the chromatin level by a recent molecular histone innovation.

PLOS Biology 17: e3000422. doi: 10.1371/journal.pbio.3000422 (*F1000 recommended*).

Histone H3 and its variants regulate gene expression but the latter is absent in most ascomycetous fungi. We identified a variant histone H3 and designated it as $H3V^{CTG}$ because of its exclusive presence in the CTG clade of ascomycetes, including *Candida albicans*, a human pathogen. *C. albicans* grows both as single yeast cells and hyphal filaments in the planktonic mode of growth. It also forms a three-dimensional biofilm structure in the host as well as on human catheter materials under suitable conditions. $H3V^{CTG}$ null (*hht1/hht1*) cells of *C. albicans* are viable but produce more robust biofilms than wild-type cells in both in vitro and in vivo conditions. A transcriptome analysis of planktonic and biofilm cells reveals that the biofilm circuitry is significantly altered in $H3V^{CTG}$ null cells. $H3V^{CTG}$ binds more efficiently to the promoters of many biofilm-related genes in the planktonic cells than during biofilm growth, whereas the binding of the core canonical histone H3 on the corresponding promoters largely remains unchanged. Furthermore, biofilm defects associated with master regulators, namely, Bcr1, Tec1, and Ndt80, are significantly rescued in cells lacking $H3V^{CTG}$. We demonstrate that co-occurrence of valine and serine at the 31st and 32nd positions in $H3V^{CTG}$, respectively, is essential for

its function. Taken together, we show that even in a unicellular organism, differential gene expression patterns are modulated by the relative occupancy of the specific histone H3 type at the chromatin level.

10. Thakur J, **Sanyal K*. (2012)** A coordinated interdependent protein circuitry stabilizes the kinetochore ensemble to protect CENP-A in the human pathogenic yeast *Candida albicans*.

PLOS Genet. 8: e1002661

This study reveals an unusual and novel interdependent assembly of the kinetochore proteins that form the macromolecular structure on the centromere DNA in *Candida albicans*. It was also demonstrated that in the absence of any of the kinetochore proteins, the CENP-A gets delocalized and degraded by the ubiquitin-mediated proteasomal machinery in the cell. All the experiments and analysis of data were done in the nominee's laboratory.

*corresponding/co-corresponding author