Finding drivers of the evolution of testis expressed genes using scRNA sequencing of Great Apes' testes

Abstract:
A large number of genes are expressed specifically in testis. We find that these genes are more likely than other genes to display evidence of positive selection in Great Apes. Out of 865 genes, we have found 211 genes with site-specific and/or branch-specific dN/dS>1 (LRT, P-value<0.05). If a testis expressed gene is found on the X-chromosome it is more likely to show signs of positive selection than if on an autosome. The X chromosome has previously been reported to harbor megabase-sized regions of low diversity that show overlap between Great Apes and are devoid of archaic admixture. Recurrent selective sweeps and hybrid incompatibilities in these regions could be explained by meiotic drive. As a result of asynapsis in pachytene, sex chromosomes are silenced during meiosis –a process known as meiotic sex chromosome inactivation (MSCI), and appear to continue repressed after meiosis, where spermatids show overall low expression after histone to protamine transition. Nevertheless, some genes have been shown to escape this repressed state in spermatids. Cytoplasmic bridges between haploid spermatids are thought to permit the exchange of transcripts, thus avoiding possible transmission ratio distortions of X and Y chromosomes. However, as recently reported, a considerable proportion of mammalian genes seem to be specific to their haploid cell's genotype. Alternatively, spermatids store transcripts necessary for later stages in the chromatoid body (CB). This ribonucleoprotein particle is also known to accumulate pachytene piRNAs, which are devoid of repetitive elements and account for ~95% of all piRNAs in adult mouse testis. Recently, these piRNAs have been reported to target specific transcripts for translation in late stages of spermatogenesis in mice. Using single-nuclei RNA sequencing of human and chimpanzee testis, we study the transcriptome of X and Y-bearing haploid spermatids, thus avoiding noise from exchanged transcripts between the products of meiosis. However, CBs appear to be attached to single nuclei, making these cells phenotypically diploid. We develop a classifier that controls for these nuisance factors, successfully classifying a moderate proportion (<50%) of X and Y-bearing spermatids. Moreover, we improve on this classification, developing a classifier that not only shows consistency with the initial classification, but is able to infer the genotype of the entire spermatids' population. Using this last classification, we identify genes that are differently expressed in X and Y-bearing spermatids, thus being putative initial candidates for meiotic drive systems. A great number of these genes show signs of positive selection. Furthermore, we determine genetic regulatory networks using the expression of transcription factors and target genes. However, gene expression in these cells seems to be only one facet of the complex transcriptome of spermatids, with stored and exchanged transcripts having a major role in post-transcriptional regulation. Using both nuclei and whole cell scRNA-Seq for the same individual, we aim to differentiate between these three sets of transcripts, thus characterizing possible differences between function and genome of X and Y bearing haploid cells. We expect to shed light into the hypothesized intragenomic conflict in the evolutionary history of Great Apes.