FHA sortseq figure packet

2021-12-01

Contents

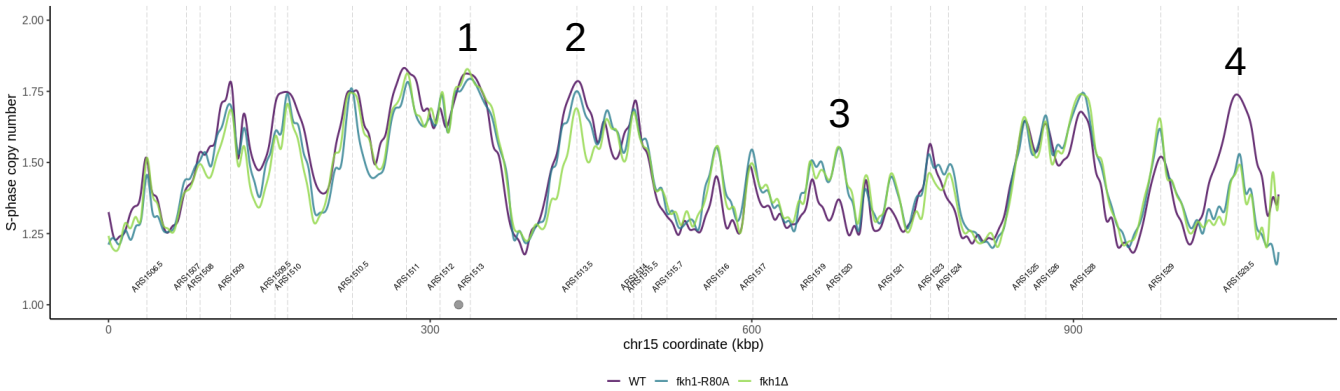
Take-homes

1. Two mutant *fkh1* alleles, *fkh1R80A* and *fkh1*Δ, affect chromosomal replication in *direct* measurement of DNA copy number in unperturbed yeast cultures.
2. The replication signatures of *fkh1R80A* and *fkh1*Δ S-phases are similar in terms of extent / amplitude of change with some important differences.
3. Single locus analysis of changes in replication signal from wildtype to either of the two mutant *fkh1* alleles suggest that sensitivity is not overwhelming shared. That is, *fkh1R80A* and *fkh1*Δ SORT sensitivity are largely distinct phenomena for orgins.
4. There is a link between *fkh1R80A* SORT sensitivity and *fkh1R80A*-mediated changes in ORC-origin binding.

Questions remaining

1. What are the smoothed signals associated with all telomeres and centromeres? Is replication at telomeres distinctly sensitive to *fkh1R80A* yet resistant to being improved in *fkh1*Δ?
2. What if any overlap exists between origins with *improved* replication?
3. What if any other features define origin cohorts? That is are there sequence and chromatin features that are enriched in any or some of the cohorts identified as fkh1R80A sensitive in either the SORT-seq or ORC ChIP-seq experiments
4. *ARS1529.5* is one of the most fkh1R80A-sensitive origins in both the SORTseq and the ORC ChIP-seq experiments. *ARS1529.5* is uniquely a positive-chromatin origin (robust ORC-origin binding yet lacking robust ORC-origin DNA affinity) *and* a Orc1BAH-dependent origin (requires the presence of the Orc1BAH for full ORC-origin binding). Is there a link between *ARS1529.5*'s fkh1R80A sensitivity and its dependence on the Orc1BAH? If we were to relax our definition of positive-chromatin ORC-origin binding, would we pick up similar Orc1BAH-depenent and positive-chromatin that are similarly *fkh1R80A* sensitive for ORC-origin binding and origin replication?

Figure one



Draft from 2021_11_29...Rmd

Figure 1: Smoothed S-phase copy numbers of chromosome 15 in wildtype (purple), *fkh1R80A* (blue-green), and *fkh1*Δ (green) cells. Positions of confirmed origins are indicated by dashed, vertical lines in gray and labeled with the origin name. The centromere is indicated by a circle on the graph. Four regions of interest are labeled with numbers on the graph. Number 1 highlights a centromere proximal initiation event that overlaps well among the three genotypes. Region 2 shows a region with initiation having differential sensitivity to the mutant *fkh1* alleles. Region 3 is an example where the two mutant alleles *increase* replication signal. Region 4 is clear example of a thoroughly *fkh1R80A* and *fkh1*Δ sensitive origin.

Figure Two

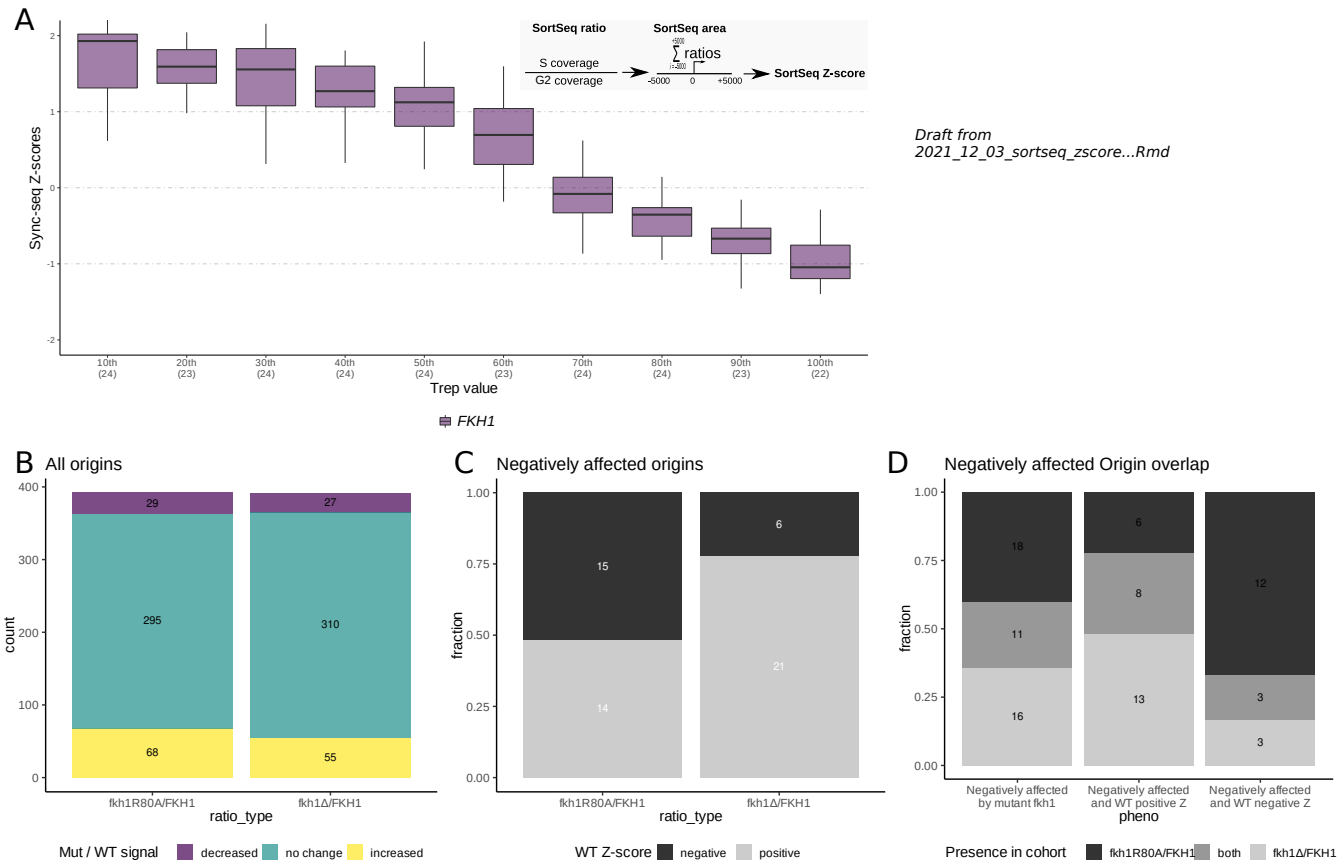
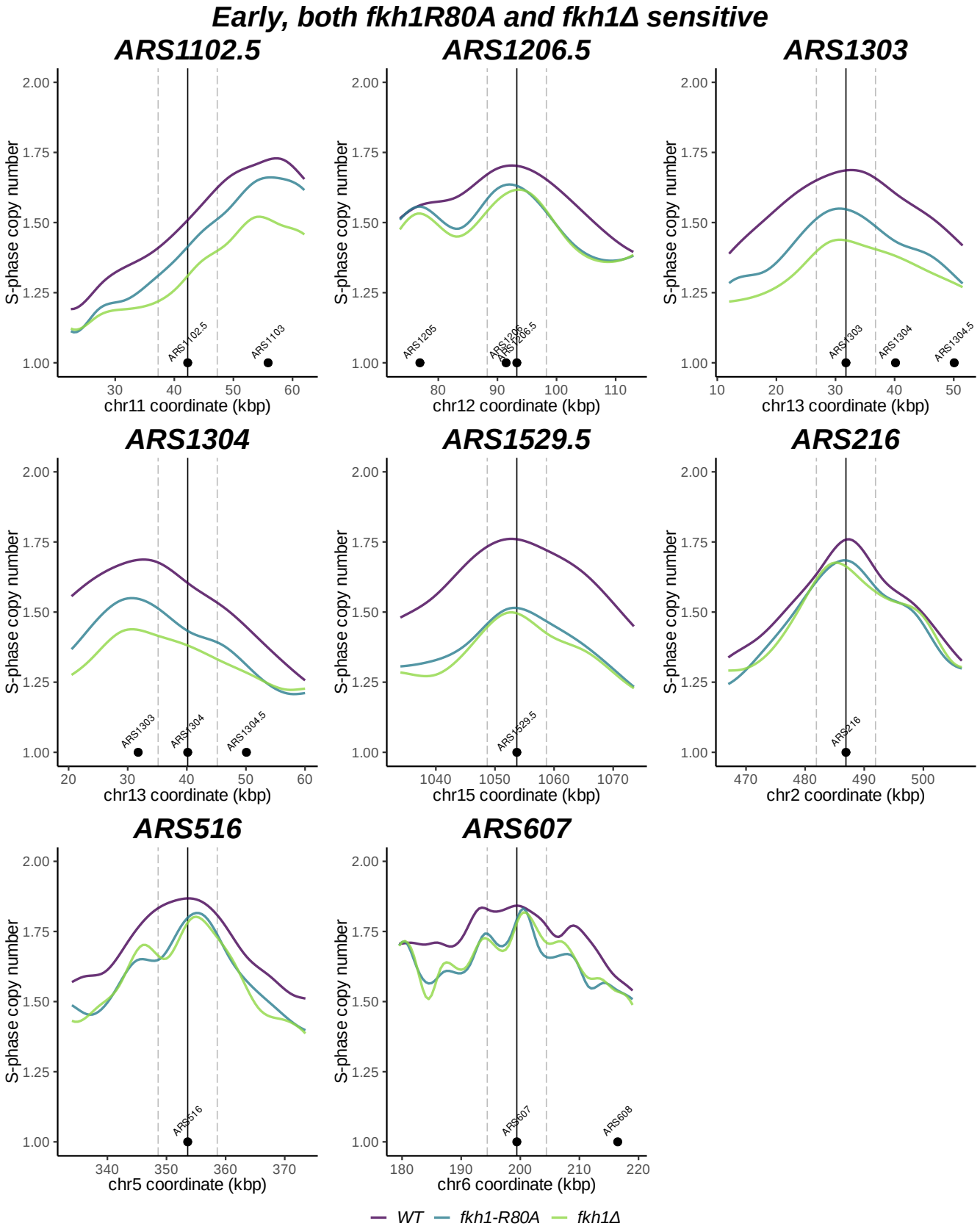


Figure 2: Z-score analysis of SORT-seq signal within discrete origin replicons. **A. The wildtype Z-scores measured in this experiment recapitulated previously defined yeast timing program (Yabuki et al).** Origins were placed into one of ten cohorts based on published Trep score, with the 10th cohort having the earliest Trep values and cohort 100th having the latest Trep values. Distributions of Z-scores of S-phase copy numbers measured in wildtype cells were summarized as box-and-whisker plots for each of the cohorts. **B. Relatively few origins show modest to severe sensitivity to either of the mutant *fkh1* alleles in unperturbed yeast.** Number of origins defined by mutant *fkh1* / *FKH1* Z-score ratio. Origins with decreased ratios (purple) were defined as ratios ≤ 0.7, those with ratios ≥ 1.3 as increased (yellow), and all ratios between (green) as no change. **C. Cohorts differ in wildtype replication.** Fraction of origins negatively affected by either the *fkh1R80A* or *fkh1*Δ allele with positive or negative Z-scores in wildtype cells. **D. Origins cohorts negatively affected by either *fkh1R80A* or *fkh1*Δ are generally distinct.** Origin overlap analysis of all negatively-affected by either of both of the mutant *fkh1* alleles. Negatively affected cohorts are further parsed by the wildtype Z-score.

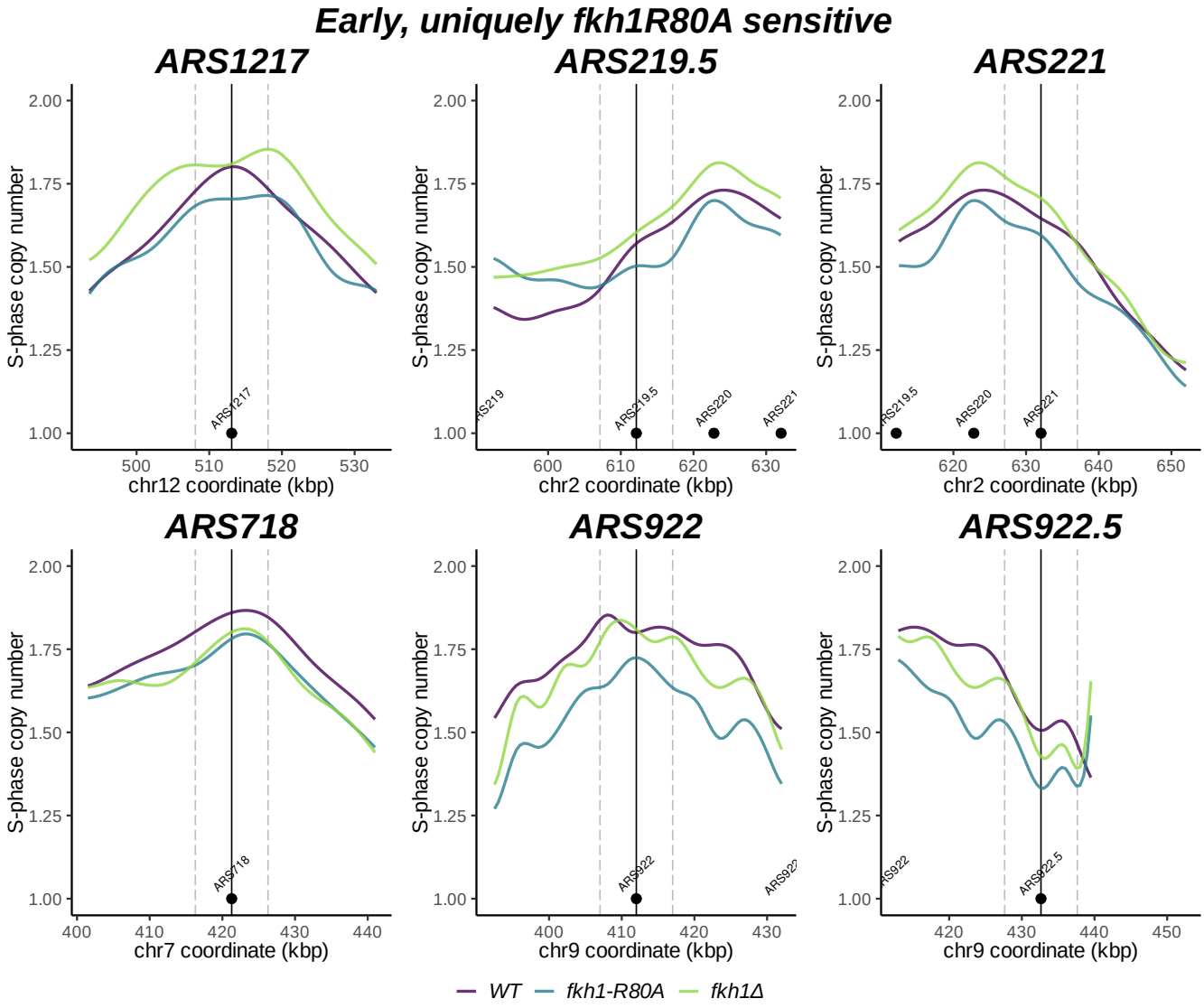
Supplemental Three

Purpose: QC to ensure that our Z-score analysis is not qualitatively different than what is observed in the chromosomal scans. The start and ending positions of the Z-score replicon fragment are indicated by vertical, dashed lines.

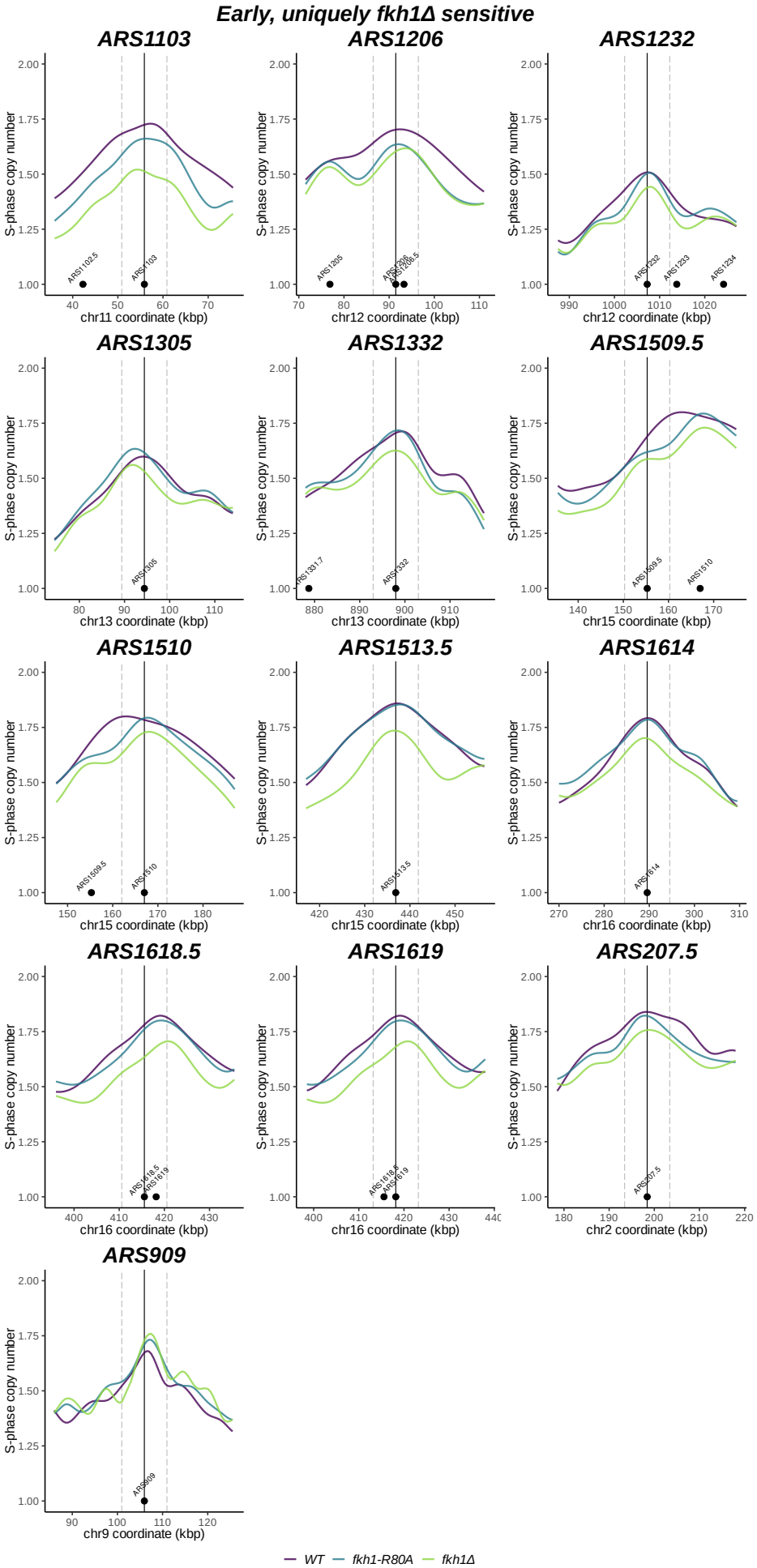
SORTseq smoothed scans of early origins SORT-sensitive to both *fkh1R80A* and *fkh1Δ*



SORTseq smoothed scans of early origins SORT-sensitive only to *fkh1R80A*



SORTseq smoothed scans of early origins SORT-sensitive only to *fkh1Δ*



SORTseq smoothed scans of late origins SORT-sensitive only to *fkh1R80A*

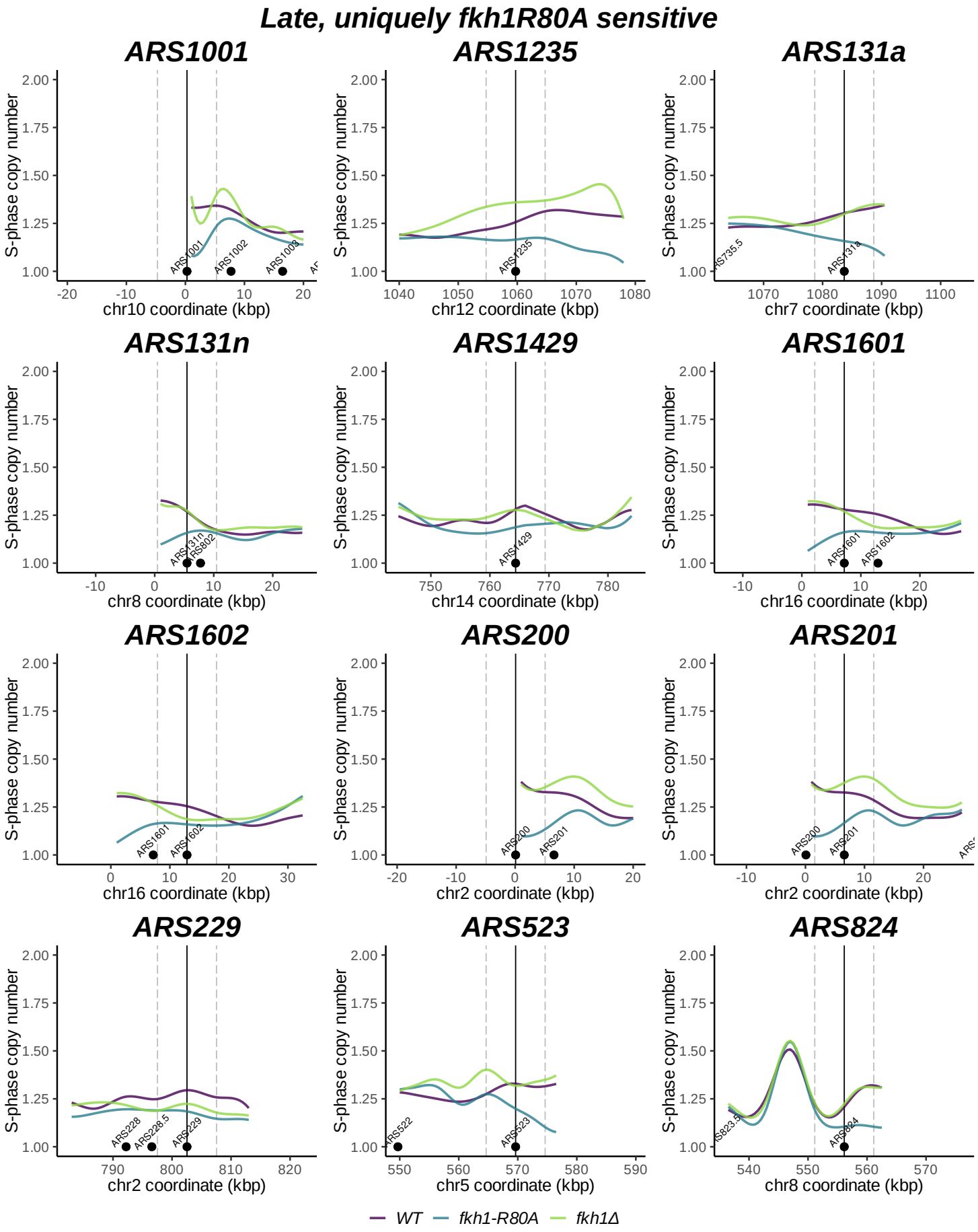


Figure Three

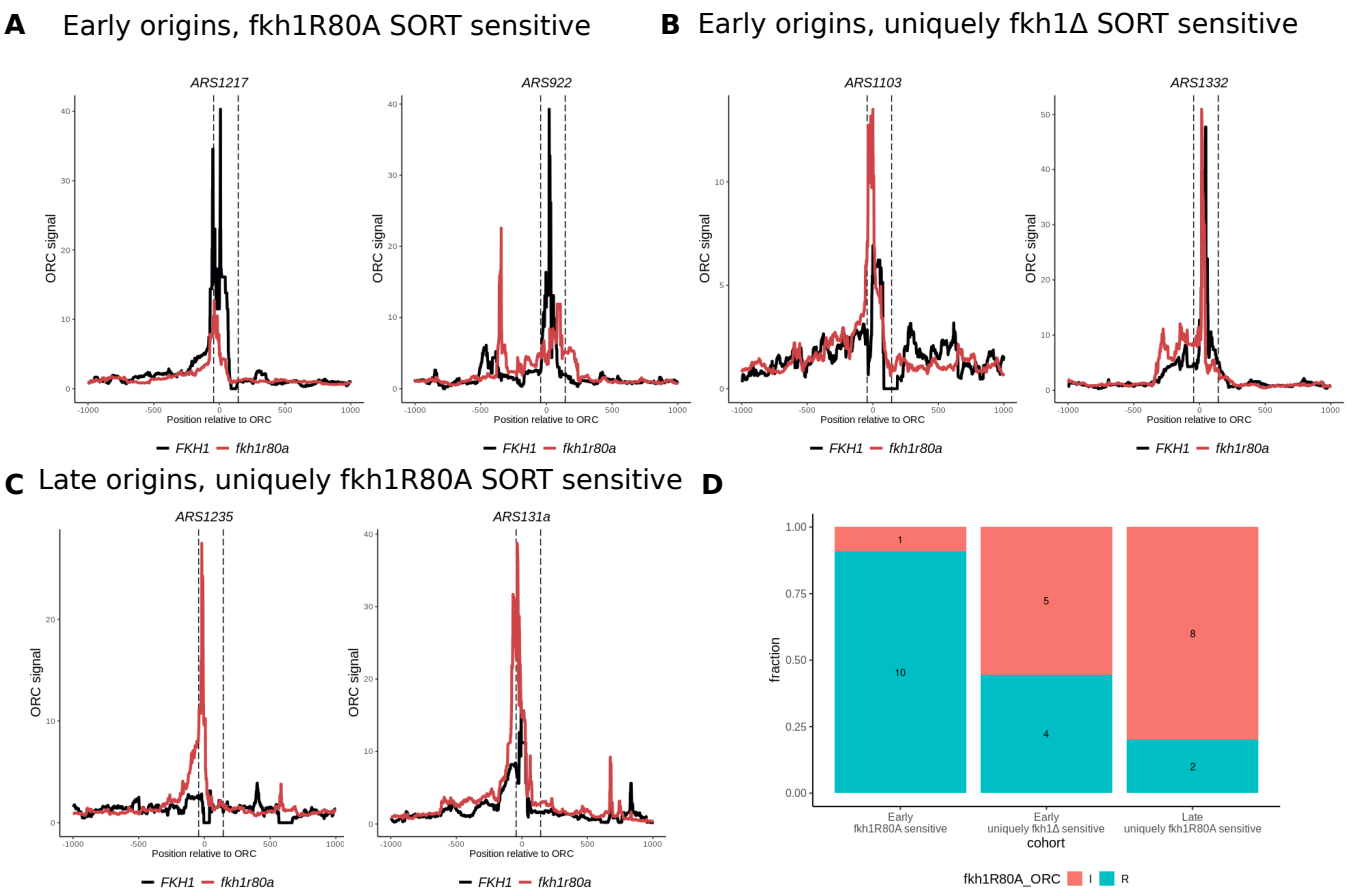


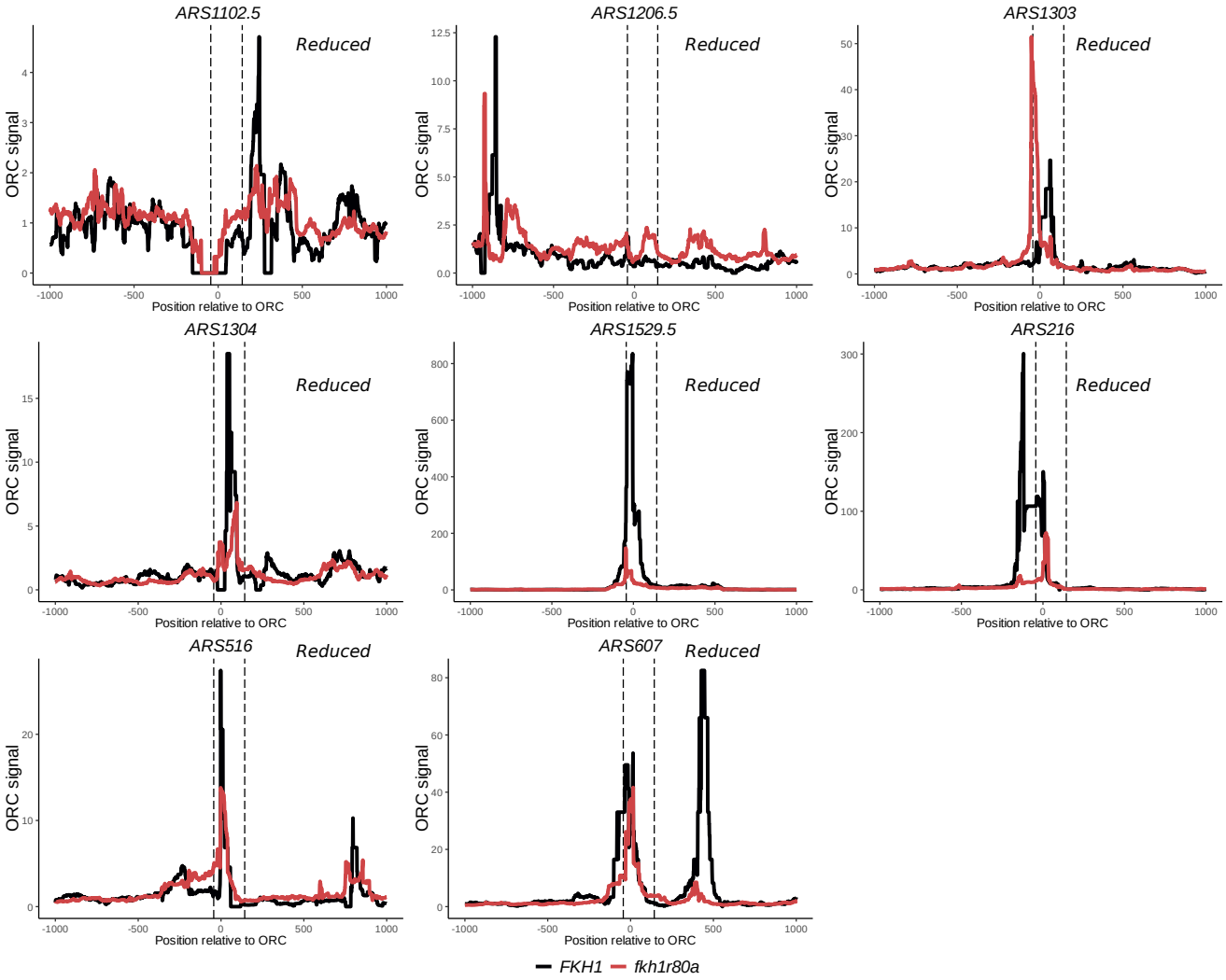
Figure 3: Fkh1FHA-dependent ORC signal at origins defined by the SORT-seq experiments. **A.** Example ORC signals associated with two early origins (i.e. positive WT Z-scores) that were *fkh1R80A* SORT sensitive as defined by Z-score ratio ≤ 0.7. **B.** Example ORC signals associated with two early origins *uniquely fkh1*Δ SORT sensitive, that is Z-score ratio ≤ 0.7 in *fkh1*Δ but > 0.7 in *fkh1R80A* cells. **C.** Example ORC signals of two late origins (i.e. negative WT Z-sccores) *uniquely fkh1R80A* SORT sensitive. **D. SORT-defined cohorts show differential enrichments of Fkh1FHA-dependent ORC binding.** All origins from the three categories in panels A-C were defined by the change in ORC binding from wildtype to *fkh1R80A* cells as either being increased (I - red), decreased (R - teal), or no-change / ambigious (not plotted) by *fkh1R80A*. The fractions and numbers of each defined subcohort are plotted as stacked barplots. The cohort 'Early fkh1R80A sensitive' contains all early origins negatively affected by either fkh1R80A alone or both fkh1R80A and fkh1Δ.

Supplemental Four

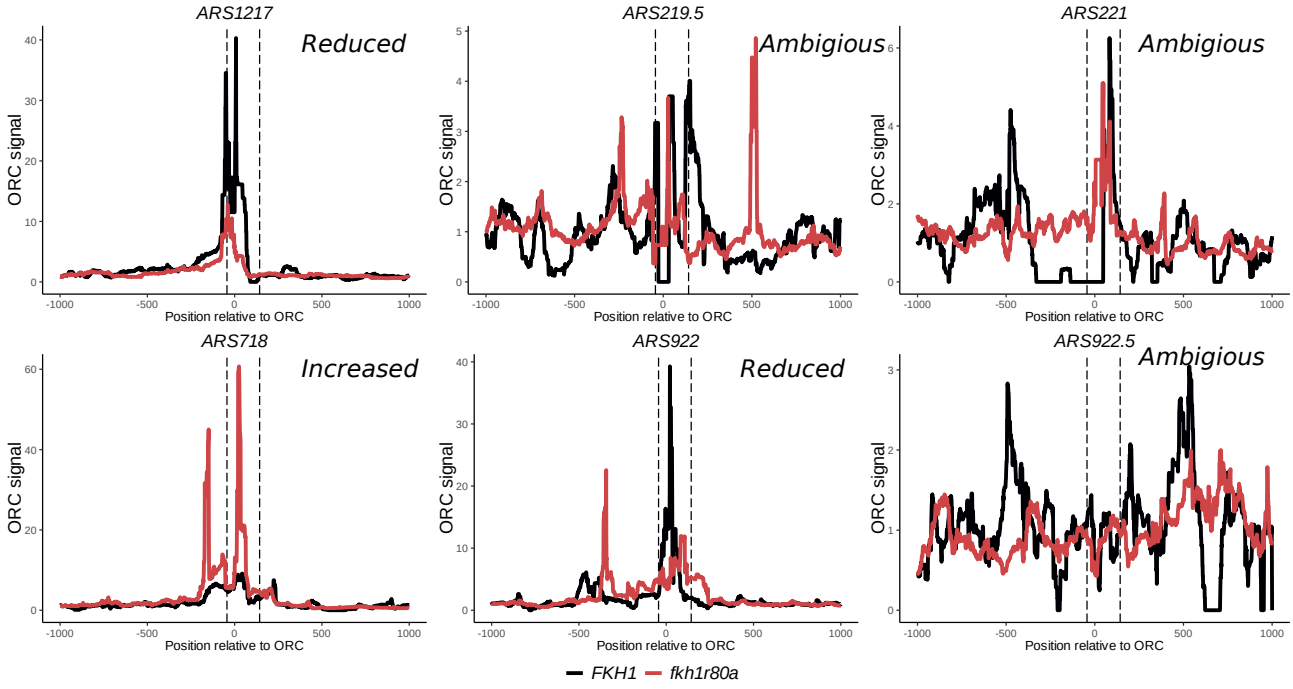
The following figures contain the ORC signals in *fkh1R80A* and *FKH1* cells for each of the origins within the SORTseq defined cohorts in this study. For each origin, the change in ORC binding in *fkh1R80A* was qualitatively assessed and

ORC scans of early origins SORT-sensitive to both *fkh1R80A* and *fkh1*Δ

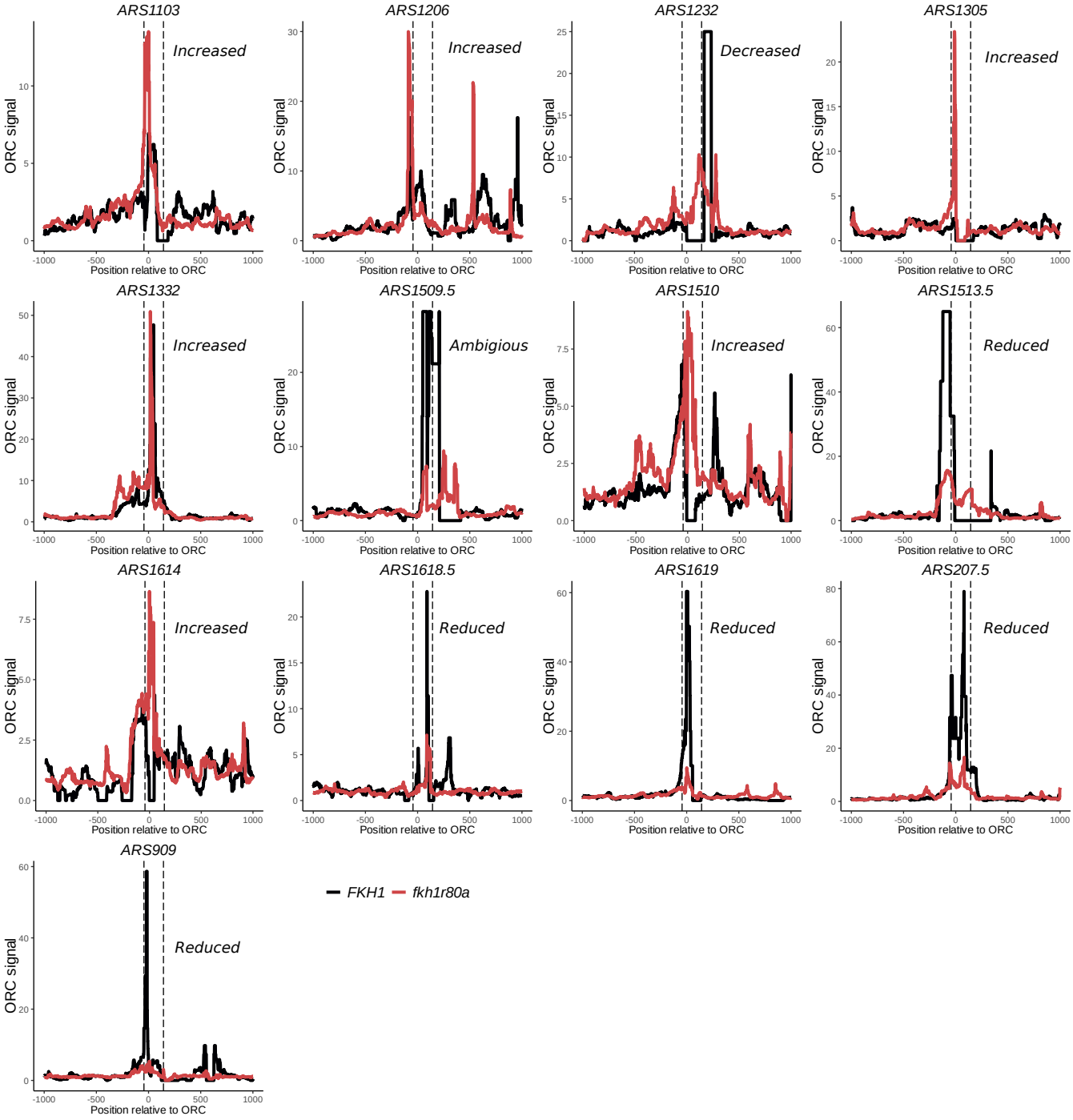
Early, fkh1R80A & fkh1Δ SORT sensitive



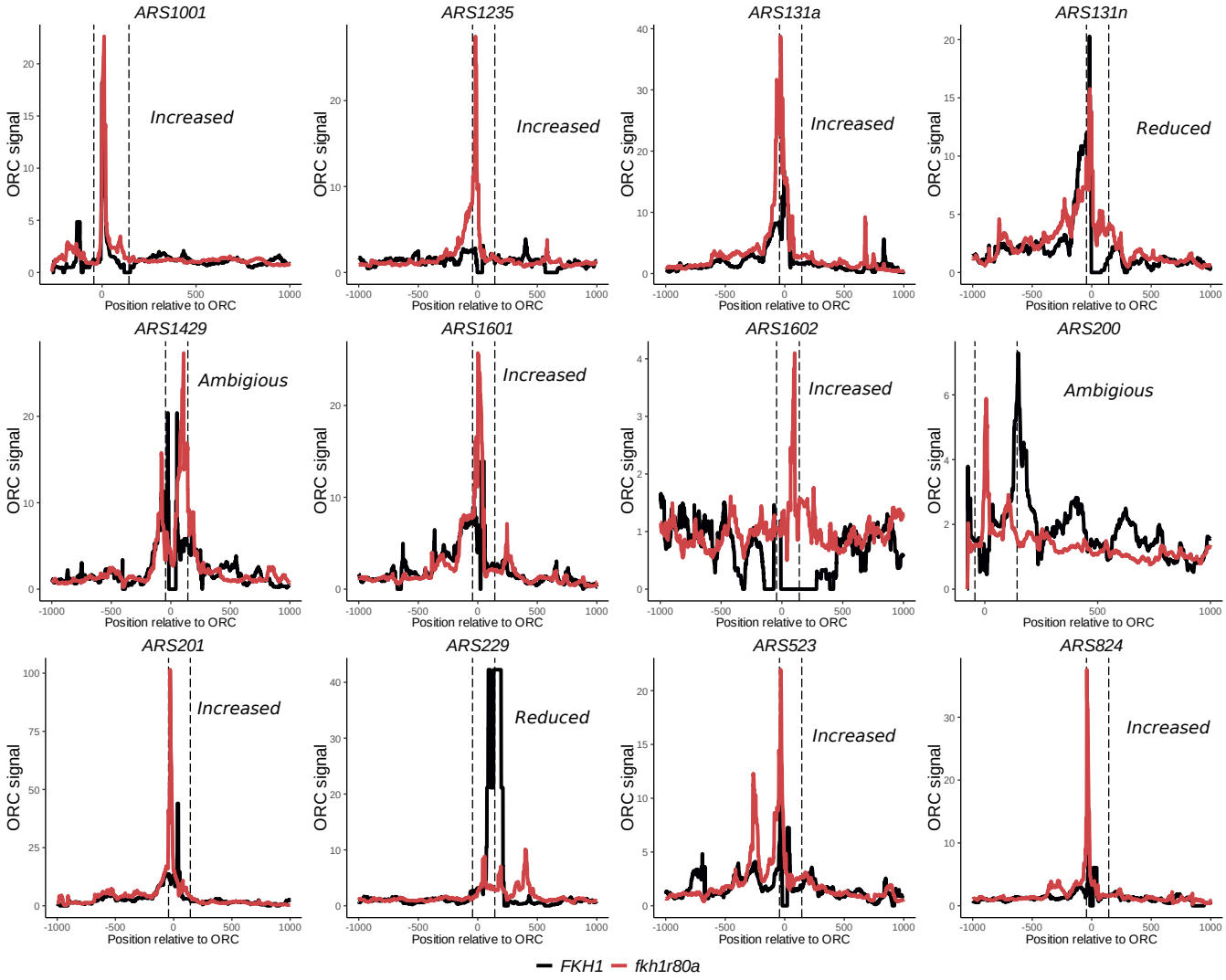
Early, uniquely fkh1R80A SORT sensitive



Early, uniquely fkh1Δ SORT sensitive



ORC scans of late origins, SORT-sensitive *only* to *fkh1R80A*



Prewriting

In this figure packet, we will address the following three, nested questions. First, to what extent do two mutant *fkh1* alleles, *fkh1-R80A* and *fkh1*Δ, impact replication at yeast origins? To answer, we will use a combination of qualitative and quantitative data to measure the number and extent of change in replication signal from wildtype to either of the two mutant alleles. Subquestions that can be addressed are do the two different alleles differ in terms of number of origins affected, either positively or negatively and to extent of change, that is how different is the mutant replication signal relative to the wildtype signal.

To this end we will present the chromosomal scans from our experiment, including the scans associated with origins measured as *fkh1-R80A* sensitive in a previous copy number - ddPCR experiment.
A) Chromosomal scan of chromosome 15. Highlight the regions around *ARS1529.5*. Importantly, there is a disconnect between *fkh1-R80A* and *fkh1*Δ replication downstream of *ARS1529.5*. Signal in *fkh1*Δ cells indicative of occassional initiation, yet signal is suppressed in *fkh1-R80A* cells. This may hint at differences in specificity? Observation raises the question of how many times do *fkh1-R80A* and *fkh1*Δ quantitatively differ?
B) Selected scans associated with *fkh1-R80A* origins tested by ddPCR in the NAR.

The second question addressed is *which* origins are affected by either of the two mutant alleles and how much overlap exists between the collection of affected origins. To this end, we will classify origins into cohorts reflecting either positive or negative replication signal change. We can thus plot the overlap between positively or negatively affected *fkh1-R80A* or *fkh1*Δ origins.

To this end, we will introduce our Z-score method.
A) Z-scores in this experiment recapitulate the timing program

The third question probed is what to any features distinguish *fkh1-R80A*, *fkh1*Δ, and *fkh1-R80A / Δ* positively affected, negatively affected relative to those origins that do not change from wildtype to mutant *fkh1*.

Random notes:/
Few origins show large changes in SortSeq Z-score. ARS1529.5 is one of the few origins that is *severly* affected by *fkh1-R80A*. An angle of this paper is to take the *ARS1529.5*-like origins and characterize them. ARS1529.5 is both *fkh1-R80A* and *fkh1*Δ sensitive. How many of the other origins are likewise doubly sensitive? What about origins that are the opposite of *ARS1529.5* - origins that show improved origin function in *fkh1-R80A*? How many and how many are improved under both mutant genotypes?

We want to focus on only those origins that should the greatest change in replication signal because we ae dealing with but one replicate in our experiment. As such, our confidence is confined to large effects.

There are more origins that are negatively affected by *fkh1*Δ relative to *fkh1-R80A*. Thus, *fkh1*Δ is a more severe allele relative to *fkh1-R80A*. Yet, there are instances of origins that are negatively affected by *fkh1-R80A* but not by *fkh1*Δ.

2021-12-03

Today, we are exploring how the origins and non-origin loci of interest that we identified yesterday appear in both the smoothed SORTSeq ratios and the ORC ChIP seq. That is we have a collection of origins that are negatively affected by either the *fkh1-R80A*, the *fkh1$ * allele, or in far fewer cases (this is an important result), origins negatively affected by* both* mutant *fkh1* alleles. At a minimum, we want to double check that the smoothed scans confirm the sensitivity to either or both of the *fkh1* alleles as measured by changes in Z-scores calculated for 10 kbp replicons. Secondly, we will take the collection of late replicating non-confirmed loci, comprising likely and dubious origins as well as non-origin intergenic sequence, that become early in either or both the mutant *fkh1* alleles. This will be a qualitative analysis. Do the loci *look* right? If the loci look *right* in the smoothed sortseq, we can extend a per-locus analysis to our published ORC ChIP-seq.

We have a nice looking template for graphing the smoothed S-phase copy numbers.

There apears to be a trend where telomeric origins are particularly sensitive to *fkh1R80A* but not to *fkh1*Δ.