

## Datasets used in this study

**Table T1 represents the alignment of the yeast WGD duplicates, the orthologs from *L. kluyveri*, *Z. rouxii* and the inferred ancestral sequences. The data about phosphorylation (fields **p** and **p.1**) comes from Dataset 1 (filtered, see methods) and (fields **p.moses** and **p.1.moses**) 2 (literature curated).** For each residue we provide information about the phosphorylation status (data from large scale studies and from low-throughput experiments), information on structure of the protein (ordered or disordered) and the predicted ancestral residue. Also indicated is the protein kinase associated with the phosphosite based on the PWM from Mok *et al* (Mok *et al*, 2010).

**Table T2 represents the alignment of the yeast WGD duplicates, the orthologs from *L. kluyveri*, *Z. rouxii* and the inferred ancestral sequences. The data about phosphorylation (fields **p** and **p.1**) comes from Dataset 3 (unfiltered, see methods) and (fields **p.moses** and **p.1.moses**) 2 (literature curated).** For each residue we provide information about the phosphorylation status (data from large scale studies and from low-throughput experiments), information on structure of the protein (ordered or disordered) and the predicted ancestral residue.

**Table T3 represents the alignment of the yeast WGD duplicates, the orthologs from *L. kluyveri*, *Z. rouxii* and *K. thermotolerans* and the inferred ancestral sequences. The data about phosphorylation (fields **p** and **p.1**) comes from Dataset 1 (filtered, see methods) and (fields **p.moses** and **p.1.moses**) 2 (literature curated).** For each residue we provide information about the phosphorylation status (data from large scale studies and from low-throughput experiments), information on structure of the protein (ordered or disordered) and the predicted ancestral residue.

## Legend

abs_pos	absolute position
rel_pos	name of the sequence, position of the residue along the sequence ( <i>S. cerevisiae</i> , paralog1)
db	dataset(s) where the phosphosite has been identified if the residue is a phosphorylation site; NA otherwise
Y1	amino acid residue
dis/ord	order/disorder status of this position (the symbol “.” means that this position along the sequence belongs to an ordered region; the symbol “*” means that this position along the sequence belongs to a disordered region)
p	phosphorylation status of the residue
rel_pos	name of the sequence, position of the residue along the sequence ( <i>S. cerevisiae</i> , paralog2)

db	dataset(s) where the phosphosite has been identified if the residue is a phosphorylation site; NA otherwise
Y2	amino acid residue
dis/ord	order/disorder status of this position (the symbol “.” means that this position along the sequence belongs to an ordered region; the symbol “*” means that this position along the sequence belongs to a disordered region)
p	phosphorylation status of the residue
anc_wag	reconstructed ancestral sequence using the Wag substitution matrix
anc_dayh of	reconstructed ancestral sequence using the Dayhoff substitution matrix f
rel_pos	name of the sequence, position of the residue along the sequence ( <i>Z. rouxii</i> ortholog)
ZYRO	amino acid residue
rel_pos	name of the sequence, position of the residue along the sequence ( <i>L. kluyveri</i> ortholog)
SAKL	amino acid residue
p.3	phosphorylation status of the residue
p.moses	phosphorylation status of the residue according to the data by Ba and Moses (paralog1)
p.1.moses	phosphorylation status of the residue according to the data by Ba and Moses (paralog2)
kin.mok	associations phosphosite-kinase generated by using the Position Weight Matrices (PWM) empirically derived by Mok <i>et al</i> (paralog1)
kin.1.mok	associations phosphosite-kinase generated by using the Position Weight Matrices (PWM) empirically derived by Mok <i>et al</i> (paralog1)