



Turnover of protein phosphorylation evolving under stabilizing selection

Christian R. Landry^{1,2,3}*, Luca Freschi^{1,2,3}, Taraneh Zarin⁴ and Alan M. Moses^{4,5,6}

¹ Département de Biologie, Université Laval, Québec, QC, Canada

² Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Québec, QC, Canada

³ Network for Research on Protein Function, Structure, and Engineering (PROTEO), Université Laval, Québec, QC, Canada

⁴ Department of Cell and Systems Biology, University of Toronto, Toronto, ON, Canada

⁵ Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON, Canada

⁶ Center for Analysis of Genome Evolution and Function, University of Toronto, Toronto, ON, Canada

Edited by:

Andreas Zanzoni, Inserm TAGC
UMR1090, France
Allegra Via, Sapienza University, Italy

Reviewed by:

Manuela Helmer-Citterich, University
of Rome Tor Vergata, Italy
Pedro Beltrao, European Molecular
Biology Laboratory, European
Bioinformatics Institute, UK

*Correspondence:

Christian R. Landry, Département de
Biologie, Université Laval, 1030
Avenue de la Médecine, Québec, QC
G1V 0A6, Canada
e-mail: christian.landry@bio.ulaval.ca

Most proteins are regulated by posttranslational modifications and changes in these modifications contribute to evolutionary changes as well as to human diseases. Phosphorylation of serines, threonines, and tyrosines are the most common modifications identified to date in eukaryotic proteomes. While the mode of action and the function of most phosphorylation sites remain unknown, functional studies have shown that phosphorylation affects protein stability, localization and ability to interact. Two broad modes of action have been described for protein phosphorylation. The first mode corresponds to the canonical and qualitative view whereby single phosphorylation sites act as molecular switches that either turn on or off specific protein functions through direct or allosteric effects. The second mode is more akin to a rheostat than a switch. In this case, a group of phosphorylation sites in a given protein region contributes collectively to the modification of the protein, irrespective of the precise position of individual sites, through an aggregate property. Here we discuss these two types of regulation and examine how they affect the rate and patterns of protein phosphorylation evolution. We describe how the evolution of clusters of phosphorylation sites can be studied under the framework of complex traits evolution and stabilizing selection.

Keywords: protein phosphorylation, evolutionary turnover, molecular switches, molecular rheostats, protein evolution, molecular evolution, cell signaling

INTRODUCTION

The rate of discovery of new protein forms is increasing with the growing sensitivity of biochemical, analytical and bioinformatics tools (Smith et al., 2013). We now contemplate the idea that a large fraction of biological diversity originates in mechanisms that regulate protein expression and functions posttranscriptionally and posttranslationally. Among the major sources of posttranslational regulation and cellular complexity are posttranslational modifications (PTMs; Jensen, 2006), which are additions of peptides, chemical groups or other complex molecules to proteins that modify their activity, stability, degradation, localization, and ability to interact (Sprang et al., 1988; Madeo et al., 1998; Vazquez et al., 2000; Khmelinskii et al., 2009). Hundreds of such modifications have been reported in the literature and some of these appear to be playing dominant roles, at least in terms of occurrence. Protein phosphorylation on serines, threonines, and tyrosines dominates by an order of magnitude the number of experimental PTMs recorded in common databases (Khoury et al., 2011; Lu et al., 2013). This domination likely derives from biases resulting from the long historical interest for protein phosphorylation (Cohen, 1982), from the more advanced state of experimental identification methods for these PTMs and also for biological reasons, for instance because of the large number of protein kinases in eukaryotic genomes that can perform the necessary enzymatic

reactions. The impact of protein phosphorylation on the regulation and deregulation of protein functions in human diseases such as Alzheimer's disease (Grundke-Iqbal et al., 1986; Alonso et al., 1996), and the numerous gains and losses of phosphorylation in cancer cells (Reimand and Bader, 2013; Reimand et al., 2013) suggest that it plays a major role in proteome regulation and in complex cellular phenotypes. For these reasons, there has been much interest in the recent years for understanding how these PTMs evolve (Moses and Landry, 2010). However, the study of phosphorylation site evolution has met several difficulties that derive from the complex mapping between PTMs and protein functions.

MOLECULAR MECHANISMS OF PROTEIN REGULATION BY PHOSPHORYLATION

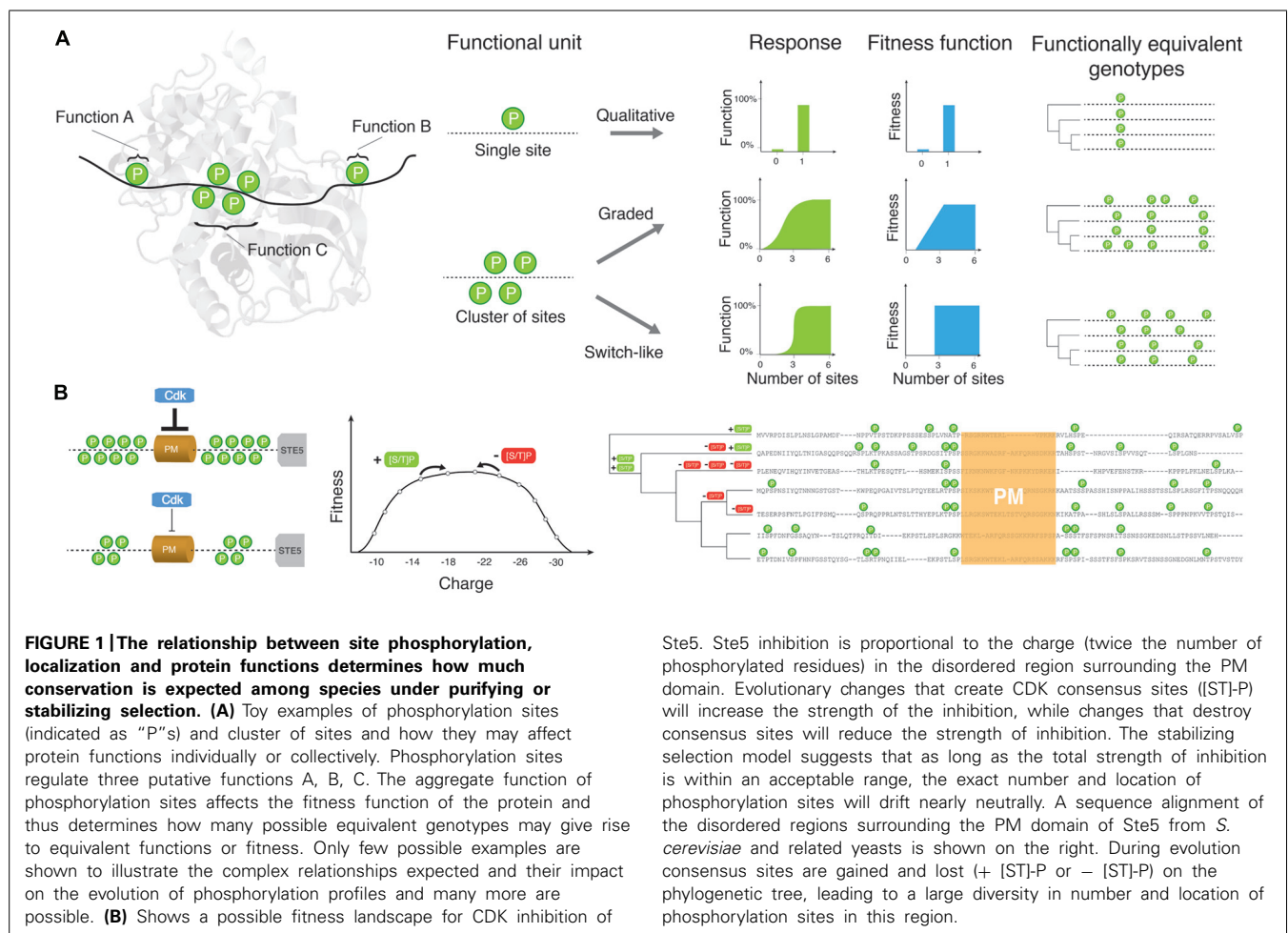
Phosphorylation sites play roles in allosteric and orthosteric regulation of proteins (Nussinov et al., 2012). Allosteric regulation acts through long-distances and involves a conformational change of the protein, while orthosteric regulation occurs at the active site of an enzyme or at the interface between a protein and another molecule. The study of protein phosphorylation has been historically centered on the role of individual phosphorylation sites. Indeed, a single phosphorylation site may have dramatic effects in regulating protein functions. For instance, the phosphorylation

of tyrosine 527 on the protein kinase and oncogene Src inactivates the protein through the interaction of this modified residue with its SH2 domain, which closes the kinase into an inactive conformation (Frame, 2002). Because single phosphorylation sites in specific cases play key roles in protein regulation, mutations at these sites may have complex organismal phenotypes. For instance, mutation of Ser47 on the *Drosophila* circadian clock protein PER modifies its interactions with other circadian proteins and lengthens adult locomotor activity from 24 to 31 h (Blau, 2008; Chiu et al., 2008). It has become clear that the simplistic view of one phosphorylation site – one function cannot be generalized to all phosphosites. Proteins are often multi phosphorylated and the different sites may affect each other's functions (Cohen, 2000).

Multisite phosphorylation is the premise of more complex types of regulation, for instance concerted regulation and modular regulation (reviewed in Salazar and Hofer, 2007). In the first case, all phosphorylation sites on a protein regulate one or more protein functions in a concerted manner (Figure 1). In the second case, groups of phosphorylation sites are organized into modules of multiple sites found in a short distance in a particular domain or disordered region of a protein and each cluster regulates a particular and independent function. Each of the mechanisms described above has its own complexity in terms of effects on the

protein and the dynamics of activation, and may thus affect the evolution of these sites (Figure 1). It is important to note that these mechanisms are not mutually exclusive. They are typically combined with various types of logic to encode patterns of substrate protein activity. For example, phosphorylation of some sites might “prime” (or increase the probability of phosphorylation) of other sites within the same region through the binding of other proteins that enhance the efficiency of the kinase (Cohen, 2000; Koivomagi et al., 2013; McGrath et al., 2013). In other cases, several sequential regulatory steps have been shown to be required for protein function (Yuan et al., 2002).

Multisite phosphorylation is associated with complex dynamic responses in signaling cascades. The number of sites being phosphorylated and their chronological order of modification can lead to graded or switch-like responses (Figure 1), depending on several parameters such as enzyme and substrate concentrations, binding parameters and the kinase/phosphatase processivity (Salazar and Hofer, 2007). One example of graded response involves the enhancement of p53 binding to CREB-binding protein (CBP). The p53 transactivation domain mediates the interaction with CBP and HDM2. Phosphorylation of Thr18 in this domain regulates the qualitative binding (on/off) of HDM2. However, the binding to CBP is regulated in a graded manner, with phosphorylation



events contributing additively to the binding energy of p53 to CBP (Lee et al., 2010). Accordingly, it is the sum of the effect that provides an appropriate function to the sites. Another recent example of aggregate effect comes from the circadian rhythm protein FRQ in *Neurospora*, which has more than 100 phosphorylation sites. It was recently demonstrated that the phosphorylation of FRQ is progressive and leads to a buildup of charge in one region of the protein that eventually leads to its degradation (Menet and Rosbash, 2011; Querfurth et al., 2011). A similar mechanism was found to control membrane association of Ste5 in the yeast mating pathway (Serber and Ferrell, 2007). The feature of interest here is that each phosphorylation site does not have a precise role but rather contributes to an aggregate property. The aggregate property was demonstrated experimentally in only few cases but clusters of phosphorylation sites are so abundant in proteomes that a large fraction of proteins could be regulated this way. Large-scale studies have indeed found that phosphorylation sites tend to localize in dense clusters of serines and threonines, which often tend to be phosphorylated by the same kinases (Moses et al., 2007a; Schweiger and Linial, 2010), supporting the hypothesis that a fraction of phosphorylation sites could regulate protein functions in a concerted way rather than acting as individual switches. However, these large-scale datasets need to be interpreted with caution. As it is the case for single phosphorylation sites (see below), clusters of phosphorylation sites could also appear as a result of non-functional phosphorylation. The tendency of phosphorylation sites to cluster could in this case result from the fact that regions accessible to protein kinases are not randomly distributed and form cluster of accessible sites that are susceptible to these phosphorylation events.

EVOLUTION OF PHOSPHORYLATION SITES

The observation that a large fraction of proteins are posttranslationally modified led to the hypothesis that changes in PTM may contribute to a large fraction of phenotypic variation within species and divergence among species (Moses and Landry, 2010). Many studies have thus examined the conservation and divergence of PTMs, particularly phosphorylation sites. Given the diversity of molecular functions that can be regulated by phosphorylation sites either individually, or when they are found in a multi-site phosphorylated protein, it is expected that their patterns of evolution would also be diverse. For example, in the classical paradigm where a single phosphorylation event leads to a conformation change that alters enzyme activity (Cohen, 1982), one predicts that the phosphorylation site and its position are conserved by purifying selection, as long as the regulation of that enzyme activity is important. Indeed, some well-characterized examples are highly conserved sites over evolution (Rittenhouse et al., 1986; Landry et al., 2009; Tan et al., 2009; Beltrao et al., 2012). However, surprisingly, other well-characterized phosphorylation sites that regulate enzyme activity are not conserved (Hwang and Fletcher, 1986; Landry et al., 2009; Boulais et al., 2010; Nguyen Ba and Moses, 2010; Freschi et al., 2011, 2014; Levy et al., 2012) suggesting that regulation of those enzymes is not critical, or that other mechanisms exist that make phosphorylation sites and their positions unessential. While the phenotypic consequences of the changes in individual sites among species is still mostly unknown

– with a few spectacular exceptions (Lynch et al., 2011) – the estimates of the relative rate of phosphorylation site evolution based on large samples led to some challenging conclusions. Some studies concluded that phosphorylation sites are generally under strong evolutionary constraints, i.e., that they evolve much slower than non-modified residues (e.g., Gray and Kumar, 2011), while others estimated that the average constraint imposed on proteins by their phosphorylation is relatively weak (e.g., Landry et al., 2009; studies reviewed in Levy et al., 2012).

Beyond the fact that variation in the different methods and datasets may contribute to some of the disagreements in estimating phosphorylation site conservation, much of the debate comes from the fact that some authors focused their attention on cases where purifying selection is strong whereas others focused on cases where there is no or little purifying selection. On the first hand, in cases where phosphorylation is known to play an important role in regulating the protein, phosphorylation sites are often strongly conserved as predicted (Landry et al., 2009; Nguyen Ba and Moses, 2010; Beltrao et al., 2012). On the other hand, the more challenging observation is that large numbers of uncharacterized sites are poorly conserved among species. There are several possible reasons why sites would evolve quickly. The first is that databases reporting large-scale data on phosphorylation may be populated with a significant fraction of false-positive identifications, i.e., sites that are not actually phosphorylated in cells. Although this limitation contributes little to our understanding of the evolution of phosphorylation sites, it will be an important challenge to be addressed by investigators developing instruments and analytical tools. Another possible scenario is that the rate of evolution is elevated because a significant fraction of sites are species-specific, i.e., they have evolved only recently by directional selection and tests of conservation among species reject the hypothesis that they are under evolutionary constraint. While this hypothesis is of biological interest, there is currently very little data supporting this possibility (but see Jensen et al., 2006; Kim and Hahn, 2011; Lynch et al., 2011), at least not on a scale that would affect significantly the results of analyses performed on thousands of phosphorylation sites.

One other possible mechanism that would explain why many phosphorylation sites evolve as non-modified residues would be that they provide no function to the protein, i.e., they result from non-functional encounters between kinases and their substrates (Lienhard, 2008; Landry et al., 2009). Because kinases are highly processive enzymes and their recognition motifs are highly degenerate (Ubersax and Ferrell, 2007), many collisions between kinases and proteins may result in the inconsequential phosphorylation of residues. This cause is obviously extremely difficult to demonstrate because it is almost impossible to show that some trait or molecular feature has no function, as it is impossible to test all possible parameters that could reveal its role experimentally. However, there are lines of evidence that support this model. For instance, bona fide functional phosphorylation sites are more conserved than the ones for which no evidence is available (Landry et al., 2009; Nguyen Ba and Moses, 2010; Freschi et al., 2014). In addition, the relative rate of phosphorylation site conservation decreases with protein abundance and increases with the stoichiometry of phosphorylation, another observation consistent

with a model by which the prevalence of non-functional phosphorylation increases with protein abundance, due to the increased probability of encounters between kinases and other proteins (Levy et al., 2012). However, this observation is also consistent with a higher rate of false positive phosphorylation sites being identified in mass spectrometry studies for high abundance proteins. We note that two of these explanations (false positives, non-functional encounters) posit no or limited biological significance to poorly conserved phosphorylation sites, whereas species specific directional selection would indicate great biological significance for these sites. A final possible explanation for the elevated rate of evolution of phosphorylation site is the rapid turnover of sites caused by the weak constraint on their localization on the protein. This is the case that we consider further here, where rapidly evolving phosphorylation sites do have biological functions, but function is not strongly dependent on each individual site, as discussed above in Section “Molecular Mechanisms of Protein Regulation by Phosphorylation.”

STABILIZING SELECTION ACTING ON CLUSTERS OF PHOSPHORYLATION SITES

The turnover of phosphorylation sites found in clusters of sites is an appealing hypothesis to explain at least a fraction of rapidly evolving phosphorylation sites (Landry et al., 2009). Indeed, a large fraction of phosphorylation sites are overrepresented in disordered regions of proteins where they occur in clusters of several juxtaposed sites (Moses et al., 2007a; Landry et al., 2009; Schweiger and Linial, 2010) and could act as groups rather than individual sites as described in Section “Molecular Mechanisms of Protein Regulation by Phosphorylation” (Figure 1). Sites could be gained and lost within a given region with limited effect on the overall function, but contributing to an increased rate of evolution if one considers sites individually. Different clusters could also play similar roles and diverging species could lose and gain these clusters without any significant change in function (Drury and Diffley, 2009). This model, which represents a previously undescribed form of “stabilizing selection” (Burger and Lande, 1994; Charlesworth, 2013a,b), suggests that clusters of sites rather than the sites themselves are the functional unit. This model leads to several predictions that can be tested using existing data: (i) protein functions regulated by phosphorylation should be preserved despite the fact that phosphorylation sites move positions; (ii) phosphorylated residues that occur in clusters should evolve faster because any change in position would be identified as an evolutionary modification; (iii) the number rather than the position of phosphorylation sites within clusters should be preserved over evolution; (iv) sites that are gained and lost between orthologs proteins should be gained and lost within these clusters. While there are no clear experimental data supporting most of these predictions, there are several observations that are consistent with them.

Observations supporting the first prediction come from proteins involved in DNA replication that are regulated by cyclin dependent kinases (CDKs). These proteins show conserved regulation in animals and fungi but diverge in the position and number of clusters of CDK phosphorylation sites, suggesting that functions may be preserved despite changes in phosphorylation

patterns (Moses et al., 2007b). Additional evidence comes from the large-scale study of CDK-dependent sites in budding yeasts where it was suggested that phosphorylation sites present in the model species but absent in closely related species could in fact be present in these other species but at other positions (Holt et al., 2009). The second prediction has been tested by (Nguyen Ba and Moses, 2010) who examined whether gains and losses of phosphorylation sites were more permissive in proteins with a high number of phosphorylation sites but no significant evidence was found against the null hypothesis. Evidence supporting predictions (iii) and (iv) comes from the analysis of yeast and mammalian phosphoproteomes. In the first case, it was shown that despite the low rate of phosphorylation site conservation (here the phosphorylation status of the proteins was compared between species or between paralogs), the actual number of phosphorylation sites tended to be maintained over evolution between homologous proteins (Freschi et al., 2011) or group of proteins (Beltrao et al., 2009). Similarly, orthologs of bona fide targets of protein kinases were found to contain conserved clusters of phosphorylation sites, often with little amino acid sequence conservation (Lai et al., 2012). Finally, it was shown that birth and death of phosphorylation sites in proteins tended to be clustered in space (Freschi et al., 2011, 2014), supporting the last prediction of the model.

More work will be needed along these lines to test whether the stabilizing selection model can apply to clusters of phosphorylation sites. Under this model, a phosphorylation cluster would be considered as a quantitative trait where selection does not act on a particular site but rather on the number or density of sites. In this model, each site represents a single locus that contributes to a “complex trait.” Deviation from the trait optimum is associated with a fitness cost that would increase or decrease following a given function (see, e.g., Charlesworth, 2013a) when phosphorylation sites are added or removed by mutations in this region (Figure 1). As mutation-selection-drift equilibrium is reached after a long divergence time, there will be a large diversity of “genotypes” that encode the trait. Even if the optimal trait value is the same in all species, if the number of phosphorylation sites (loci) that contribute to the trait is large, a minimal level of conservation will be observed for the actual phosphorylation sites. For example, CDK-mediated inhibition of Ste5 membrane localization is mediated by 8 phosphorylation sites in *S. cerevisiae* (Figure 1; Serber and Ferrell, 2007; Strickfaden et al., 2007). The strength of the inhibition is proportional to the net charge of the Ste5 PM domain after phosphorylation by CDKs. In evolutionary terms, the net charge of the PM domain is a quantitative trait whose value simply doubles the sum of the CDK phosphorylation sites in the PM domain. If a mutation occurs at a phosphorylation site, and renders it unphosphorylatable, the net charge of the PM domain will decrease by 2. On the other hand, if a mutation creates a new phosphorylation site (by adding a serine or threonine followed by a proline) in this region, the value of the trait will increase by 2. Assuming that there are several values of the trait that are compatible with functional inhibition by CDK, the PM domain will be free to drift in sequence space through any “genotype” that has enough phosphorylation sites to retain function. At mutation-selection-drift equilibrium, a large number of genotypes will coexist (Figure 1),

corresponding to different positions and numbers of phosphorylation sites in the PM domain. This stabilizing selection model has been applied to explain the rapid evolution of transcription factor binding sites in highly conserved developmental enhancers in *Drosophila* (Ludwig et al., 2000). In one recent simulation study, clusters of transcription factor binding sites were shown to evolve spontaneously simply because there are a much larger number of genotypes with many binding sites that encode the trait than “simple” genotypes with few binding sites (He et al., 2012). Whether this type of model can explain the origin of phosphorylation site clusters is currently a promising open research question.

CONCLUSION

Given the current knowledge on the function of clusters of phosphorylation sites, there is a need for the development of new evolutionary models or adaptation of existing ones that take into account the fuzziness of phosphorylation site position and density. Although to our knowledge the stabilizing selection model has not been explicitly applied to the evolution of phosphorylation sites, the distribution of kinase site consensus matches in the absence of selection has been calculated (Lai et al., 2012). This represents a neutral null hypothesis for the evolution of phosphorylation site clusters. Deviations from this model were exploited to predict substrates for several protein kinases because orthologs of bona fide substrates contain more matches to the phosphorylation site consensus than expected in the absence of selection. Knowledge on kinase recognition motifs can thus be exploited to examine the neutral evolution of phosphorylation sites and thus to derive null hypothesis regarding their evolution.

The study of protein phosphorylation will also need more experimental studies on the function of cluster of phosphorylation sites in order to learn the general principles by which their aggregate properties emerge from their combination. This will allow for instance to estimate the relationship between the number of sites, their density and positions, and protein function and organismal fitness. Most evolutionary studies performed so far, with few exceptions (Landry et al., 2009), rely on comparative data among species and not on within species polymorphisms. Detailed analysis of within species variation coupled with functional studies could help estimate the distribution of genotypes under selection-mutation-drift equilibrium for given clusters of sites. Above all, the adaptation of more complex models to the evolution of PTMs will provide a better global picture of the evolutionary forces acting on phosphoproteomes. At the same time, these models will have the potential to contribute to biochemical studies whereby evolutionary observations will guide experimental investigators in studying the code of phosphorylation site regulation.

ACKNOWLEDGMENTS

We thank Jukka-Pekka Verta, Marie Filteau, Guillaume Diss, Samuel Rochette, and Alexandre Dubé for comments on the manuscript. This work was supported by Canadian Institute of Health Research (CIHR) Grants GMX-191597 and GMX-299432, a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery grant and a Human Frontier Science Program grant to Christian R. Landry. Christian R. Landry is a CIHR

New Investigator. Luca Freschi was supported by fellowships from the Fonds de Recherche du Québec – Nature et Technologies (FRQ-NT) and the Quebec Research Network on Protein Function, Structure and Engineering (PROTEO). Alan M. Moses was supported by grants from CIHR (grant MOP-119579) and NSERC.

AUTHOR CONTRIBUTIONS

Christian R. Landry drafted the manuscript with contributions from Alan M. Moses. Luca Freschi contributed concepts, discussion and to manuscript edition. Taraneh Zarin and Luca Freschi contributed the figure.

REFERENCES

- Alonso, D. F., Farias, E. F., Famulari, A. L., Dominguez, R. O., Kohan, S., and De Lustig, E. S. (1996). Excessive urokinase-type plasminogen activator activity in the euglobulin fraction of patients with Alzheimer-type dementia. *J. Neurol. Sci.* 139, 83–88. doi: 10.1016/0022-510X(96)00036-6
- Beltrao, P., Albanese, V., Kenner, L. R., Swaney, D. L., Burlingame, A., Villen, J., et al. (2012). Systematic functional prioritization of protein posttranslational modifications. *Cell* 150, 413–425. doi: 10.1016/j.cell.2012.05.036
- Beltrao, P., Trinidad, J. C., Fiedler, D., Roguev, A., Lim, W. A., Shokat, K. M., et al. (2009). Evolution of phosphoregulation: comparison of phosphorylation patterns across yeast species. *PLoS Biol.* 7:e1000134. doi: 10.1371/journal.pbio.1000134
- Blau, J. (2008). PERSpective on PER phosphorylation. *Genes Dev.* 22, 1737–1740. doi: 10.1101/gad.1696408
- Boulais, J., Trost, M., Landry, C. R., Dieckmann, R., Levy, E. D., Soldati, T., et al. (2010). Molecular characterization of the evolution of phagosomes. *Mol. Syst. Biol.* 6, 423. doi: 10.1038/msb.2010.80
- Burger, R., and Lande, R. (1994). On the distribution of the mean and variance of a quantitative trait under mutation-selection-drift balance. *Genetics* 138, 901–912.
- Charlesworth, B. (2013a). Stabilizing selection, purifying selection, and mutational bias in finite populations. *Genetics* 194, 955–971. doi: 10.1534/genetics.113.151555
- Charlesworth, B. (2013b). Why we are not dead one hundred times over. *Evolution* 67, 3354–3361. doi: 10.1111/evo.12195
- Chiu, J. C., Vanselow, J. T., Kramer, A., and Edery, I. (2008). The phosphorylation occupancy of an atypical SLIMB-binding site on PERIOD that is phosphorylated by DOUBLETIME controls the pace of the clock. *Genes Dev.* 22, 1758–1772. doi: 10.1101/gad.1682708
- Cohen, P. (1982). The role of protein phosphorylation in neural and hormonal control of cellular activity. *Nature* 296, 613–620. doi: 10.1038/296613a0
- Cohen, P. (2000). The regulation of protein function by multisite phosphorylation – a 25 year update. *Trends Biochem. Sci.* 25, 596–601. doi: 10.1016/S0968-0004(00)01712-6
- Drury, L. S., and Difley, J. F. (2009). Factors affecting the diversity of DNA replication licensing control in eukaryotes. *Curr. Biol.* 19, 530–535. doi: 10.1016/j.cub.2009.02.034
- Frame, M. C. (2002). Src in cancer: deregulation and consequences for cell behaviour. *Biochim. Biophys. Acta* 1602, 114–130. doi: 10.1016/S0304-419X(02)00040-9
- Freschi, L., Courcelles, M., Thibault, P., Michnick, S. W., and Landry, C. R. (2011). Phosphorylation network rewiring by gene duplication. *Mol. Syst. Biol.* 7, 504. doi: 10.1038/msb.2011.43
- Freschi, L., Osseni, M., and Landry, C. R. (2014). Functional divergence and evolutionary turnover in mammalian phosphoproteomes. *PLoS Genet.* 10:e1004062. doi: 10.1371/journal.pgen.1004062
- Gray, V. E., and Kumar, S. (2011). Rampant purifying selection conserves positions with posttranslational modifications in human proteins. *Mol. Biol. Evol.* 28, 1565–1568. doi: 10.1093/molbev/msr013
- Grundke-Iqbal, I., Iqbal, K., Tung, Y. C., Quinlan, M., Wisniewski, H. M., and Binder, L. I. (1986). Abnormal phosphorylation of the microtubule-associated protein tau (tau) in Alzheimer cytoskeletal pathology. *Proc. Natl. Acad. Sci. U.S.A.* 83, 4913–4917. doi: 10.1073/pnas.83.13.4913

- He, X., Duque, T. S., and Sinha, S. (2012). Evolutionary origins of transcription factor binding site clusters. *Mol. Biol. Evol.* 29, 1059–1070. doi: 10.1093/molbev/msr277
- Holt, L. J., Tuch, B. B., Villen, J., Johnson, A. D., Gygi, S. P., and Morgan, D. O. (2009). Global analysis of Cdk1 substrate phosphorylation sites provides insights into evolution. *Science* 325, 1682–1686. doi: 10.1126/Science.1172867
- Hwang, P. K., and Fletterick, R. J. (1986). Convergent and divergent evolution of regulatory sites in eukaryotic phosphorylases. *Nature* 324, 80–84. doi: 10.1038/324080a0
- Jensen, L. J., Jensen, T. S., De Lichtenberg, U., Brunak, S., and Bork, P. (2006). Co-evolution of transcriptional and post-translational cell-cycle regulation. *Nature* 443, 594–597. doi: 10.1038/nature05186
- Jensen, O. N. (2006). Interpreting the protein language using proteomics. *Nat. Rev. Mol. Cell Biol.* 7, 391–403. doi: 10.1038/Nrm1939
- Khmelnitskii, A., Roostalu, J., Roque, H., Antony, C., and Schiebel, E. (2009). Phosphorylation-dependent protein interactions at the spindle midzone mediate cell cycle regulation of spindle elongation. *Dev. Cell* 17, 244–256. doi: 10.1016/j.Devcell.2009.06.011
- Khoury, G. A., Baliban, R. C., and Floudas, C. A. (2011). Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci. Rep.* 1. doi: 10.1038/srep00090
- Kim, D. S., and Hahn, Y. (2011). Identification of novel phosphorylation modification sites in human proteins that originated after the human-chimpanzee divergence. *Bioinformatics* 27, 2494–2501. doi: 10.1093/Bioinformatics/Btr426
- Koivomagi, M., Ord, M., Iofik, A., Valk, E., Venta, R., Faustova, I., et al. (2013). Multisite phosphorylation networks as signal processors for Cdk1. *Nat. Struct. Mol. Biol.* 20, 1415–1424. doi: 10.1038/nsmb.2706
- Lai, A. C., Nguyen Ba, A. N., and Moses, A. M. (2012). Predicting kinase substrates using conservation of local motif density. *Bioinformatics* 28, 962–969. doi: 10.1093/bioinformatics/bts060
- Landry, C. R., Levy, E. D., and Michnick, S. W. (2009). Weak functional constraints on phosphoproteomes. *Trends Genet.* 25, 193–197. doi: 10.1016/j.tig.2009.03.003
- Lee, C. W., Ferreon, J. C., Ferreon, A. C., Arai, M., and Wright, P. E. (2010). Graded enhancement of p53 binding to CREB-binding protein (CBP) by multisite phosphorylation. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19290–19295. doi: 10.1073/pnas.1013078107
- Levy, E., Michnick, S., and Landry, C. R. (2012). Protein abundance is key to distinguish promiscuous from functional phosphorylation based on evolutionary information. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 2594–2606. doi: 10.1098/rstb.2012.0078
- Lienhard, G. E. (2008). Non-functional phosphorylations? *Trends Biochem. Sci.* 33, 351–352. doi: 10.1016/j.tibs.2008.05.004
- Lu, C. T., Huang, K. Y., Su, M. G., Lee, T. Y., Bretana, N. A., Chang, W. C., et al. (2013). DbPTM 3.0: an informative resource for investigating substrate site specificity and functional association of protein post-translational modifications. *Nucleic Acids Res.* 41, D295–D305. doi: 10.1093/nar/gks1229
- Ludwig, M. Z., Bergman, C., Patel, N. H., and Kreitman, M. (2000). Evidence for stabilizing selection in a eukaryotic enhancer element. *Nature* 403, 564–567. doi: 10.1038/35000615
- Lynch, V. J., May, G., and Wagner, G. P. (2011). Regulatory evolution through divergence of a phosphoswitch in the transcription factor CEBPB. *Nature* 480, 383–386. doi: 10.1038/nature10595
- Madeo, F., Schlauer, J., Zischka, H., Mecke, D., and Frohlich, K. U. (1998). Tyrosine phosphorylation regulates cell cycle-dependent nuclear localization of Cdc48p. *Mol. Biol. Cell* 9, 131–141. doi: 10.1091/mbc.9.1.131
- McGrath, D. A., Balog, E. R., Koivomagi, M., Lucena, R., Mai, M. V., Hirschi, A., et al. (2013). Cks confers specificity to phosphorylation-dependent CDK signaling pathways. *Nat. Struct. Mol. Biol.* 20, 1407–1414. doi: 10.1038/nsmb.2707
- Menet, J. S., and Rosbash, M. (2011). A new twist on clock protein phosphorylation: a conformational change leads to protein degradation. *Mol. Cell* 43, 695–697. doi: 10.1016/j.molcel.2011.08.015
- Moses, A. M., Heriche, J. K., and Durbin, R. (2007a). Clustering of phosphorylation site recognition motifs can be exploited to predict the targets of cyclin-dependent kinase. *Genome Biol.* 8, R23. doi: 10.1186/gb-2007-8-2-r23
- Moses, A. M., Liku, M. E., Li, J. J., and Durbin, R. (2007b). Regulatory evolution in proteins by turnover and lineage-specific changes of cyclin-dependent kinase consensus sites. *Proc. Natl. Acad. Sci. U.S.A.* 104, 17713–17718. doi: 10.1073/Pnas.0700997104
- Moses, A. M., and Landry, C. R. (2010). Moving from transcriptional to phospho-evolution: generalizing regulatory evolution? *Trends Genet.* 26, 462–467. doi: 10.1016/j.tig.2010.08.002
- Nguyen Ba, A. N., and Moses, A. M. (2010). Evolution of characterized phosphorylation sites in budding yeast. *Mol. Biol. Evol.* 27, 2027–2037. doi: 10.1093/molbev/msq090
- Nussinov, R., Tsai, C. J., Xin, F., and Radivojac, P. (2012). Allosteric post-translational modification codes. *Trends Biochem. Sci.* 37, 447–455. doi: 10.1016/j.tibs.2012.07.001
- Querfurth, C., Diernfellner, A. C., Gin, E., Malzahn, E., Hofer, T., and Brunner, M. (2011). Circadian conformational change of the *Neurospora* clock protein FREQUENCY triggered by clustered hyperphosphorylation of a basic domain. *Mol. Cell* 43, 713–722. doi: 10.1016/j.molcel.2011.06.033
- Reimand, J., and Bader, G. D. (2013). Systematic analysis of somatic mutations in phosphorylation signaling predicts novel cancer drivers. *Mol. Syst. Biol.* 9, 637. doi: 10.1038/msb.2012.68
- Reimand, J., Wagih, O., and Bader, G. D. (2013). The mutational landscape of phosphorylation signaling in cancer. *Sci. Rep.* 3, 2651. doi: 10.1038/srep02651
- Rittenhouse, J., Harrsch, P. B., Kim, J. N., and Marcus, F. (1986). Amino acid sequence of the phosphorylation site of yeast (*Saccharomyces cerevisiae*) fructose-1,6-bisphosphatase. *J. Biol. Chem.* 261, 3939–3943.
- Salazar, C., and Hofer, T. (2007). Versatile regulation of multisite protein phosphorylation by the order of phosphate processing and protein-protein interactions. *FEBS J.* 274, 1046–1061. doi: 10.1111/j.1742-4658.2007.05653.x
- Schweiger, R., and Linial, M. (2010). Cooperativity within proximal phosphorylation sites is revealed from large-scale proteomics data. *Biol. Direct* 5, 6. doi: 10.1186/1745-6150-5-6
- Serber, Z., and Ferrell, J. E. Jr. (2007). Tuning bulk electrostatics to regulate protein function. *Cell* 128, 441–444. doi: 10.1016/j.cell.2007.01.018
- Smith, L. M., Kelleher, N. L., Linial, M., Goodlett, D., Langridge-Smith, P., Ah Goo, Y., et al. (2013). Proteoform: a single term describing protein complexity. *Nat. Methods* 10, 186–187. doi: 10.1038/nmeth.2369
- Sprang, S. R., Acharya, K. R., Goldsmith, E. J., Stuart, D. I., Varvill, K., Fletterick, R. J., et al. (1988). Structural changes in glycogen phosphorylase induced by phosphorylation. *Nature* 336, 215–221. doi: 10.1038/336215a0
- Strickfaden, S. C., Winters, M. J., Ben-Ari, G., Lamson, R. E., Tyers, M., and Pryciak, P. M. (2007). A mechanism for cell-cycle regulation of MAP kinase signaling in a yeast differentiation pathway. *Cell* 128, 519–531. doi: 10.1016/j.cell.2006.12.032
- Tan, C. S., Bodenmiller, B., Pasculescu, A., Jovanovic, M., Hengartner, M. O., Jorgensen, C., et al. (2009). Comparative analysis reveals conserved protein phosphorylation networks implicated in multiple diseases. *Sci. Signal.* 2, ra39. doi: 10.1126/scisignal.2000316
- Ubersax, J. A., and Ferrell, J. E. Jr. (2007). Mechanisms of specificity in protein phosphorylation. *Nat. Rev. Mol. Cell Biol.* 8, 530–541. doi: 10.1038/nrm2203
- Vazquez, F., Ramaswamy, S., Nakamura, N., and Sellers, W. R. (2000). Phosphorylation of the PTEN tail regulates protein stability and function. *Mol. Cell Biol.* 20, 5010–5018. doi: 10.1128/MCB.20.14.5010-5018.2000
- Yuan, J., Eckerdt, F., Bereiter-Hahn, J., Kurunci-Csacsko, E., Kaufmann, M., and Strebhardt, K. (2002). Cooperative phosphorylation including the activity of polo-like kinase 1 regulates the subcellular localization of cyclin B1. *Oncogene* 21, 8282–8292. doi: 10.1038/sj.onc.1206011

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 05 June 2014; accepted: 08 July 2014; published online: 23 July 2014.

Citation: Landry CR, Freschi L, Zarin T and Moses AM (2014) Turnover of protein phosphorylation evolving under stabilizing selection. *Front. Genet.* 5:245. doi: 10.3389/fgene.2014.00245

This article was submitted to *Systems Biology*, a section of the journal *Frontiers in Genetics*.

Copyright © 2014 Landry, Freschi, Zarin and Moses. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.