ANNUAL REVIEWS

Annual Review of Microbiology

Past, Present, and Future of Extracytoplasmic Function σ Factors: Distribution and Regulatory Diversity of the Third Pillar of Bacterial Signal Transduction

Thorsten Mascher

General Microbiology, Technische Universität Dresden, Dresden, Germany; email: thorsten.mascher@tu-dresden.de

Annu. Rev. Microbiol. 2023. 77:625–44

First published as a Review in Advance on July 12, 2023

The *Annual Review of Microbiology* is online at micro.annualreviews.org

https://doi.org/10.1146/annurev-micro-032221-024032

Keywords

σ factor, signal transduction, stress response, gene regulation, transcription initiation, comparative genomics

Abstract

Responding to environmental cues is a prerequisite for survival in the microbial world. Extracytoplasmic function σ factors (ECFs) represent the third most abundant and by far the most diverse type of bacterial signal transduction. While archetypal ECFs are controlled by cognate anti-σ factors, comprehensive comparative genomics efforts have revealed a much higher abundance and regulatory diversity of ECF regulation than previously appreciated. They have also uncovered a diverse range of anti-σ factor–independent modes of controlling ECF activity, including fused regulatory domains and phosphorylation-dependent mechanisms. While our understanding of ECF diversity is comprehensive for well-represented and heavily studied bacterial phyla—such as *Proteobacteria*, *Firmicutes*, and Actinobacteria (phylum *Actinomycetota*)—our current knowledge about ECF-dependent signaling in the vast majority of underrepresented phyla is still far from complete. In particular, the dramatic extension of bacterial diversity in the course of metagenomic studies represents both a new challenge and an opportunity in expanding the world of ECF-dependent signal transduction.

Contents

BACTERIAL SIGNAL TRANSDUCTION: CONNECTING STIMULI WITH CELLULAR RESPONSES

In the course of evolution, bacteria have conquered all niches, adapted to virtually all conditions, and adjusted their physiology to thrive in the most adverse and complex habitats on Earth. Their ability to respond to numerous environmental cues and threats in a fast and accurate manner is both a prerequisite for survival and a stunning achievement, considering the overall simplicity and diminutiveness of microbial cells. Signal transduction describes the process of connecting an extra- or intracellular stimulus that functions as an input to an appropriate cellular response as the output. These steps require the concerted action of sensory and regulatory domains that provide specificity in signal processing and gene regulation. Three fundamental principles of bacterial signal transduction have evolved to provide the molecular mechanisms that connect a stimulus with a cellular response.

In one-component systems (1CSs), sensory and regulatory domains are directly fused on a single protein. 1CSs are the most widely distributed signaling devices and are particularly suitable for responding to intracellular cues (70). Stimulus perception by the sensory domain results in a conformational change that affects the activity of the output domain. This simplified architecture provides a direct way of connecting intracellular signals in particular to cellular responses. Accordingly, the vast majority of input domains are involved in small-molecule binding, while

the majority of output domains mediate DNA binding and hence differential gene expression (7). Nevertheless, 1CSs can also control alternative output domains, including enzymatic activities involved in protein phosphorylation (e.g., Ser/Thr kinases) or in second messenger signaling (70). While combining input and output domains on one protein leads to sterical constraints on responses to extracellular cues, membrane-spanning 1CSs—which respond to extracellular cues—have been described, and approximately 5% of all 1CSs are predicted to be membrane-anchored proteins (70). Examples include BcrR, which mediates bacitracin resistance in *Enterococcus faecalis* (13), and CadC, which is involved in the acid stress response of *Escherichia coli* (8).

The term one-component systems is less well established than the term it was originally inspired by: two-component systems (2CSs). This second most widely distributed type of bacterial signaling separates the sensory and regulatory function onto two different proteins, the sensor histidine kinase and the response regulator, respectively (20, 67). This physical separation enables such systems to more readily respond to extracellular cues by connecting a membrane-anchored sensor kinase specifically to a soluble regulator through a partner-specific phosphotransfer reaction. In response to perceiving a stimulus, the sensor kinase autophosphorylates at an invariant histidine residue and subsequently transfers the phosphoryl group to the cognate response regulator to a conserved aspartate residue. In response, the regulator will typically dimerize and thereby activate its output domain. Approximately two-thirds of all response regulators mediate differential gene expression, but alternative outputs, such as RNA binding, second messenger signaling, and protein–protein interactions, are also possible (20).

Next to 1CSs and 2CSs, extracytoplasmic function σ factors (ECFs) represent the third fundamental principle of bacterial signal transduction (64, 66). ECFs are widely distributed in the bacterial world and highly diverse with regard to the molecular mechanisms orchestrating ECF-dependent signal transduction (4, 6, 9, 34, 51). This review only briefly discusses the hallmark features and paradigms of ECF-dependent regulation, which has been thoroughly reviewed elsewhere, before embarking on its primary topic: the knowledge gained by comparative genomic analyses on the astonishing diversity of ECF-dependent regulation that goes far beyond these paradigms. This review also touches upon the limitations of such analyses as well as the potential of bioinformatic predictions for mechanistic studies of ECF-dependent gene regulation in the age of genomics. Our current understanding gained from, and the predictive power embedded in, comparative genomics opens up a world of novel ECF-dependent regulatory mechanisms—a vast landscape waiting to be explored.

THE PAST: DISCOVERY AND REGULATORY PRINCIPLES OF ECFs

σ factors are essential subunits of the RNA polymerase holoenzyme and mediate promoter recognition and hence transcription initiation (30). Two phylogenetically unrelated families of σ factors can be distinguished, σ^{54} and σ^{70} (named after the corresponding proteins from *E. coli*), of which the latter is by far the most important and diverse. The σ^{70} protein family has been subdivided into four groups on the basis of sequence conservation and domain architecture of the corresponding σ factors (23, 54). All bacterial genomes encode at least one essential σ^{70} protein. These primary (or housekeeping) σ factors are classified as group I σ^{70} proteins and harbor four conserved domains, termed regions σ_1 through σ_4 (30) (**Figure 1*a***). Their nonessential paralogs, which share an identical domain architecture and include the stationary phase σ factors, such as *E. coli* σ^S, belong to group II. Closely related alternative σ factors lacking region σ_1 are classified as group III σ^{70} proteins. These proteins are involved in mediating heat-shock responses and in controlling flagellar biosynthesis or sporulation gene expression. ECFs, or group IV σ factors, are the simplest and phylogenetically most diverse members of the σ^{70} protein family. They harbor only the crucially

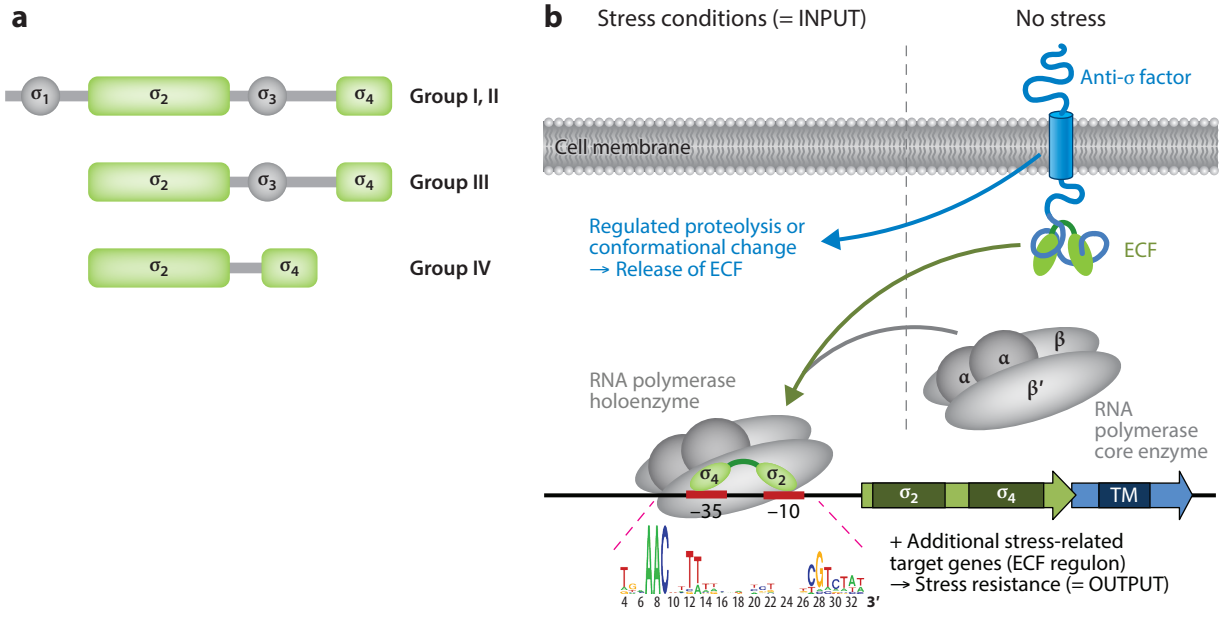

Figure 1

(*a*) Domain architecture of the different groups of σ^{70} proteins. (*b*) Overview of typical features of ECF-dependent signal transduction and gene regulation. Abbreviations: ECF, extracytoplasmic function σ factor; TM, transmembrane helix. Figure adapted with permission from Reference 50.

important and highly conserved regions, σ_2 and σ_4, which are necessary and sufficient for recognizing the bipartite promoter motifs (the −35 and −10 regions, respectively) and directing the RNA polymerase to its transcription initiation site (23, 54) (**Figure 1**). Surprisingly, and despite their very low complexity, structural studies demonstrated that the mechanism of how ECFs interact with the RNA polymerase and bind/unwind their target promoter DNA is largely identical to that of the much more complex primary σ factors (42, 43).

ECFs were initially thought to be transcriptional regulators and not σ factors, due to their small size, reduced complexity, and overall sequence diversity (29). Their discovery was the result of the concerted biochemical efforts of Mark Buttner's group and their collaboration with Michael Lonetto, a graduate student in Carol Gross's group (46). In their seminal paper, biochemical evidence from *Streptomyces cerevisiae* σ^E and its sequence similarity to a couple of related regulators, including *E. coli* RpoE and FecI; *Bacillus subtilis* SigX; *Pseudomonas aeruginosa* AlgU; and *Myxococcus xanthus* CarQ, led to their recognition as new members of the σ^{70} protein family, which were predominantly involved in responding to extracellular conditions, hence the name ECF (45).

Over the next 20 years, the ECF paradigms (see below) were comprehensively studied and a number of hallmark features of ECFs-dependent signaling were identified (26) (**Figure 1*b***):

1. ECFs are small proteins (approximately 200 amino acids) that harbor only the σ^{70} regions σ_2 and σ_4.
2. ECFs recognize alternative promoters often containing a highly conserved AAC signature in the −35 region and CGT bases in the −10 region.
3. ECFs are usually encoded in an operon with their cognate anti-σ factor (ASF). ASFs are small, predominantly membrane-anchored proteins with often very little sequence conservation (10).

4. Activation of ECFs requires their release from the inhibitory grip of the cognate ASFs, usually by regulated proteolysis or conformational changes of the ASFs.
5. The *ecf-asf* operon is subject to positive autoregulation. Therefore, the ECF-specific promoter can be found upstream of it.
6. ECFs control regulons often involved in coordinating stress responses or uptake processes.

THE PRESENT: CLASSIFICATION AND MECHANISTIC DIVERSITY OF ECFs

The age of microbial genomics began in 1995 with the first two genome sequences of *Haemophilus influenzae* (17) and *Mycoplasma genitalium* (19). A decade later, some hundred finished bacterial genomes became available for the first comprehensive comparative genomic studies of bacterial signal transduction. An analysis of 145 microbial genomes established the predominance of 1CSs over 2CSs (70). Four years later, an analysis of 369 microbial genomes (66) led to the recognition of the distribution and diversity of ECFs. This study (66) used the hallmark features of ECFs described above. Specifically, ECFs are proteins harboring the conserved σ^{70} regions σ_2 and σ_4, with a spacing of fewer than 50 amino acids in between to rule out the additional presence of region σ_3. On the basis of this domain architecture, approximately 2,700 ECF sequences were retrieved and subsequently clustered, resulting in the first ECF classification that described 67 distinct ECF groups (**Table 1**). Importantly, the ECF groups—as defined by the sequence similarity between the ECFs—mirrored the other ECF hallmark features very well. Within each ECF group, the ASFs, the target promoters, and the genomic context of the ECF gene were conserved. This congruence of different features not only allowed conserved ECF groups to be defined but also provided the predictive power to suggest completely novel mechanisms of ECF-dependent signal transduction and can serve as a reliable guideline for directing subsequent experimental studies of the molecular mechanisms orchestrating signal transduction in such novel ECF groups (for examples, see the section titled Novel Mechanisms of ECF-Dependent Regulation, Identified by Comparative Genomics).

This initial ECF classification set the stage for all subsequent efforts to expand our understanding of ECF-dependent regulation, by establishing that (*a*) ECFs are much more widely distributed in the bacterial world than initially anticipated, (*b*) they employ many novel modes of regulating their activity beyond the classical ECF–ASF interaction, and (*c*) the ECF promoter signature is more diverse than previously indicated (64, 66). While this initial study (66) established an average number of six ECFs per bacterial genome, it also highlighted an unusual bias in ECF distribution.

Table 1 Overview of ECF classification efforts

Year	Number of phyla	Genomes	ECFs	ECF groups	Reference
Initial ECF classifications					
2009	11	369	2,708	67	66
2011	4	18	21	1	21
2012	1 (planctomycetes)	8	362	8	39
2015	1 (actinomycetes)	119	2,203	18	37
2020	1 (planctomycetes)	150	5,966	30	72
ECF reclassifications and expansions					
2021	130	156,241	177,910	157	11
2023	117	4,212	16,181	43	This review

Abbreviation: ECF, extracytoplasmic function σ factor.

In contrast to 1CSs and 2CSs, which scale in numbers roughly with genome size, ECFs—for unclear reasons—are underrepresented in smaller genomes while being overrepresented in larger genomes (37, 58). Approximately 20% of all ECF groups lack a discernable ASF partner. Instead, they employ alternative mechanisms of ECF control, including fused regulatory extensions and direct or indirect phosphorylation-dependent mechanisms. And of those ECF groups harboring a conserved group-specific ASF, around 20% are soluble proteins. All of these novel features were reliably conserved in an ECF group–specific manner (66).

During the 2010s, a number of expansions of the ECF classification increased our knowledge of the distribution and diversity of ECF-dependent signal transduction. These studies basically followed the experimental approach established in the initial study but made use of the ECF group–specific sequence signatures (hidden Markov models) defined during the first ECF classification effort (66).

The Actinobacteria contain many species with complex lifestyles and extensive secondary metabolism capacities, most prominently the filamentous, multicellular streptomycetes but also important human pathogens, such as the mycobacteria and biotechnological workhorses, including *Corynebacterium* spp. This complexity is reflected in overall larger genome sizes and, correspondingly, an enrichment of signal transducing proteins. A comprehensive analysis of the signaling capacity of 119 actinobacterial species identified countless new protein architectures for 1CSs, 2CSs, and ECFs (37). With approximately 2,200 ECFs identified in these 119 species, the average number of 18 ECFs per actinobacterial genome is three times higher than the overall average initially described (66). This study (37) identified 18 new ECF groups, most of which were restricted to the Actinobacteria (**Table 1**). Four of these novel groups contain fused regulatory domains; three lack a discernable ASF and are therefore regulated by alternative mechanisms. While this study demonstrated that an increased genome sequence space allows the identification of novel ECF-dependent signaling mechanisms, it still focused on one of the best-represented bacterial phyla of the initial study (66). Two subsequent expansions (39, 72) therefore focused on an underrepresented phylum that also contains bacteria with complex lifestyles.

Planctomycetes share several unusual features, including membrane-bound compartments within their cytoplasm, differences in cell envelope architecture, the mechanism of cell division by budding, and an ability to perform endocytosis. Their enigmatic biology is reflected by large genomes and an overall richness in signaling proteins, including many novel ECF groups with proposed unique mechanisms of signal transduction. While the initial study (66) was restricted to eight genomes, a more comprehensive follow-up analysis was based on a total of 150 genomes, including 79 newly isolated, cultivated, and functionally characterized planctomycetes, leading to an in-depth understanding of these unique bacteria (39, 72). Taken together, both studies (39, 72) identified more than 6,000 ECFs, with an average of 40 ECFs per genome! Fewer than 5% could be assigned to known ECF groups. As a consequence, a total of 38 novel ECF groups were defined in this phylum alone (**Table 1**). This significant expansion of the ECF diversity within a single bacterial phylum indicates that a wide range of novel ECFs might still be hidden in the remaining >100 underrepresented bacterial phyla.

Phosphorylation-dependent ECF signaling involving Ser/Thr protein kinases seems to be particularly widespread and is found in seven of the planctomycete-specific ECF groups. Moreover, this phylum is particularly rich in ECFs with regulatory extensions, which can be found in 10 of the newly identified ECF groups. These include potentially membrane-anchored ECFs, in which long C-terminal extensions are separated from the ECF core by three transmembrane regions, indicating the presence of a spatial separation of an extracellular sensory domain and a cytoplasmic regulatory domain across the membrane. Moreover, planctomycetes harbor ECFs with N-terminal regulatory extensions (39, 72).

In 2016, a comprehensive summary described 94 ECF groups (59). Together with the 30 ECF groups identified 3 years later in another study of the planctomycetes (72), the total was around 120 ECF groups. But these phylum-specific extensions of ECF-dependent signaling diversity did not solve the increasing challenge associated with the advent of next-generation sequencing and the explosion of the microbial genome sequence landscape. By 2017, the National Center for Biotechnology Information (NCBI) database contained more than 180,000 microbial genome sequences, which required a complete overhaul of the previous ECF classification, including the development of a bioinformatics pipeline able to cope with these massive numbers. The resulting ECF reclassification not only covered a 50-fold-increased number of ECFs but also removed a number of limitations and weaknesses of the initial classification efforts, based on poorly performing group-specific sequence signatures and hence unconvincing group definitions, as noticed earlier (11, 59). This reclassification effort refined 31 of the original ECF groups and identified 22 novel groups. A total of 157 phylogenetic ECF groups were defined on the basis of conservation of ECF sequence, genetic neighborhood, target promoter motif, and common regulatory mechanism, with an average number of 10 ECFs per genome (11). Comprehensive information about ECF-dependent regulation is freely accessible through the so-called ECF Hub (see Related Resources) (11).

The next subsections describe the different mechanisms of ECF regulation. First, I introduce some of the ECF paradigms representing classical modes of ECF-dependent regulation (**Figure 2**). These include (*a*) regulated proteolysis of membrane-anchored ASFs in cell envelope stress responses (*E. coli* σ^E, *B. subtilis* σ^W), (*b*) conformational changes of soluble ASFs in oxidative stress responses (*Rhodobacter sphaeroides* σ^E, *Streptomyces coelicolor* σ^R), (*c*) ASF-mediated protein interaction cascades to control uptake processes (*E. coli* FecI), and (*d*) transcriptional regulation of ECF expression (*S. coelicolor* σ^E). Since all of these systems have been comprehensively reviewed elsewhere, I describe them only briefly here. Subsequently, I describe novel conserved modes of ECF control, initially identified and predicted by comparative genomics in the course of the ECF classification efforts. These include (*e*) σ factor mimicry, (*f*) ECF regulation by fused regulatory domains, and (*g*) σ factor phosphorylation by Ser/Thr kinases.

ECFs Controlled by Proteolysis of Membrane-Anchored ASFs

E. coli σ^E and *B. subtilis* σ^W follow the same logic for ECF activation (1, 2, 27, 63) (**Figure 2*a***): In the absence of inducing conditions, these two ECFs are tightly bound by their cognate ASFs (*E. coli* RseA and *B. subtilis* RsiW, respectively), thereby preventing their recruitment by the RNA polymerase core enzyme. In the presence of envelope stress, the ASFs are sequentially degraded in three successive steps. First, site I proteolysis of the ASF occurs in its extracytoplasmic domain through the action of the initial, membrane-anchored protease (*E. coli* DegS, *B. subtilis* PrsW), which acts as the sensor of the inducing stress signal. The release of the extracellular domain renders the ASF susceptible to intramembrane cleavage by the action of membrane-anchored RIP proteases (*E. coli* RseP, *B. subtilis* RasP), thereby releasing a soluble ASF–ECF complex into the cytoplasm. The remaining ASF fragment is then readily degraded by the cytoplasmic ClpXP protease to finally release the ECF from its inhibition (2, 25, 27, 63). A similar general layout of ASF degradation holds true for other ECFs from *B. subtilis* (σ^M, σ^V, σ^X) and *Mycobacterium tuberculosis* (σ^K, σ^L, σ^M) (28, 35, 61, 63).

ECFs Controlled by Soluble ASFs

R. sphaeroides σ^E and *S. coelicolor* σ^R are controlled by soluble ASFs (**Figure 2*b***). While σ^E-ChrR specifically responds to singlet oxygen, a reactive oxygen species, σ^R-RsrA mediates a much more complex regulation in response to a diverse range of inducing conditions, including thiol oxidants,

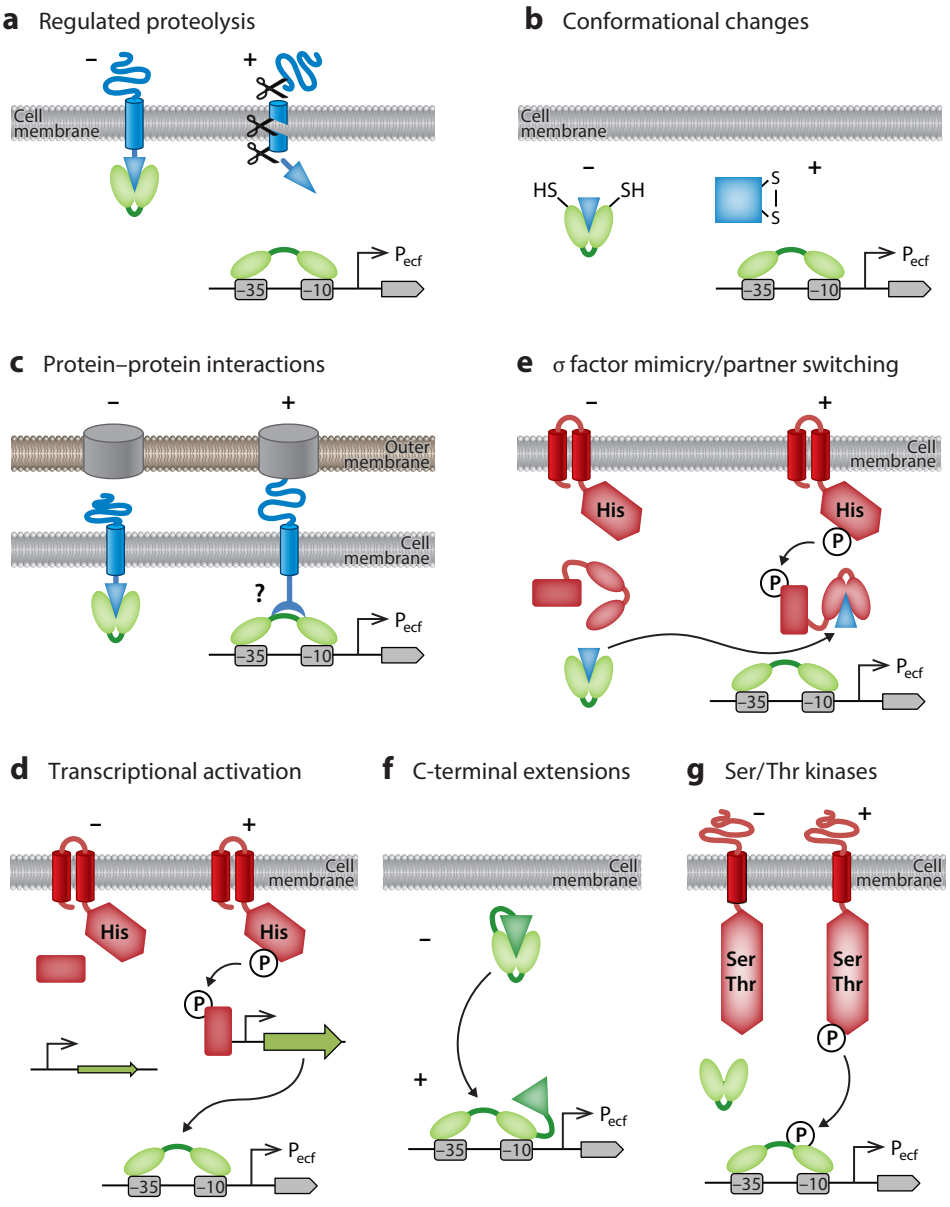

Figure 2

Different mechanisms of ECF-dependent signal transduction. (Potential) anti-σ factors are shown in blue, ECFs are shown in green, and phosphorylation-dependent signaling proteins are shown in red. A minus sign (–) indicates no stimulus, and a plus sign (+) indicates the presence of a stimulus. Abbreviations: ECF, extracytoplasmic function σ factor; His, histidine kinase; Ser/Thr, serine/threonine kinase. Figure adapted with permission from Reference 50.

alkylating electrophiles, and antibiotics interfering with translation (14, 55). While both ASFs respond to redox stress, their mechanisms of inactivation are rather different. Perception of singlet oxygen promotes ChrR proteolysis, thereby releasing σ^E to induce gene expression and counteract oxidative stress (14). In contrast, inactivation of RsrA by redox stress leads to a conformational

change based on a disulfide bridge formation between two cysteine residues. This conformational change of RsrA ultimately leads to σ^R release (55).

Despite a rather low degree of sequence conservation of the ASFs, a structurally conserved anti-σ domain (ASD), derived from the crystal structures of *E. coli* RseA, *M. tuberculosis* RskA, *B. subtilis* RsiW, and *R. sphaeroides* ChrR, has been identified in 50% of both soluble and membrane-anchored ASFs (10). Remarkably, and despite a high degree of structural similarity, the interactions of these ASDs with their cognate ECFs differ significantly (63).

ECFs Controlled Through Protein–Protein Interactions via Cell Surface Signaling

The FecI–FecR system of *E. coli* regulates Fe^{3+} uptake in the presence of the chelator citrate (5, 47, 52). This process is mediated by an intricate series of protein–protein interactions that involve the $(Fe^{3+}\text{-citrate})_2$-specific outer membrane porin FecA and its cytoplasmic membrane partner TonB (**Figure 2*c***). Their interaction triggers activation of the ECF FecI through the ASF FecR. The FecI–FecR pair is unique for several reasons. First, the membrane-anchored ASF FecR, although showing the usual protein topology, not only exhibits its expected negative function on FecI activity in the absence of $(Fe^{3+}\text{-citrate})_2$ but also is required for full FecI activity in its presence, potentially by inducing FecI to bind to the RNA polymerase core enzyme. Second, FecI seems to be unstable in the absence of FecR, probably because of proteolytic degradation. Lastly, the *fecIR* operon is not subject to positive autoregulation, in contrast to most ECFs (5, 47, 52).

Transcriptional Control of ECF Activity

Despite its pivotal role in the discovery of ECFs (45), σ^E of *S. coelicolor* is atypical in that its activity is not regulated by an ASF. Instead, σ^E is controlled by transcription of its structural *sigE* gene through a 2CS, CseBC (69) (**Figure 2*d***). CseBC-σ^E controls the cell envelope stress response of *S. coelicolor* when challenged with cell wall antibiotics (45). σ^E represents an early example of the diversity of ECF control by alternative mechanisms and demonstrates that evolution has freely combined modules from different signaling principles to generate an ever-increasing number of mechanisms underlying how bacteria connect signals with responses. Both assumptions were strongly supported by the first comparative genomic analyses of the ECF protein family, which unraveled the wide distribution and regulatory diversity of these alternative σ factors in the microbial world and helped establish them as the third most abundant mechanism of bacterial signal transduction (66).

Novel Mechanisms of ECF-Dependent Regulation, Identified by Comparative Genomics

The true value of the ECF classification is its predictive power in identifying the molecular ingredients and hence the potential mechanisms of ECF-dependent signal transduction and gene regulation, especially for completely novel ECF groups that deviate from known blueprints. The sequence similarity among ECFs allowed their identification in microbial genomes and determined their initial clustering. Within such conserved groups, additional ECF hallmark features are usually well conserved:

1. Positive autoregulation enables the identification of group-specific ECF target promoters upstream of ECF-encoding genes.
2. A group-specific promoter signature then allows the prediction of putative target genes and hence the potential physiological role of such ECFs.

3. Cotranscription of ECF- and ASF-encoding genes helps identify the ASF, which is particularly helpful because of their overall poor degree of sequence conservation.
4. Extended genomic context conservation within ECF groups enables the identification of regulatory functions or target genes, based on their cooccurrence with ECF genes.

Together, these predictive features can significantly streamline efforts to unravel the signaling mechanisms of novel ECF groups. Three such novel mechanisms of controlling ECF function that have been characterized (in part), inspired by the ECF classification, are described below: (*a*) The ECF-mediated general stress response of *Alphaproteobacteria* (group ECF15), which combines 2CS- and ECF-dependent signaling; (*b*) ECFs that lack ASFs but instead use fused regulatory domains to control ECF activity (e.g., groups ECF41, ECF42, and ECF238); and (*c*) ECFs that are controlled through phosphorylation by Ser/Thr kinases (e.g., group ECF43).

The Best of Both Worlds: Combining ECFs and Two-Component Systems to Generate a Partner-Switching Module Based on σ Factor Mimicry

Most bacteria can mount a global and transient general stress response to overcome severe but nonspecific stress conditions, orchestrated by alternative σ factors. The gram-negative and gram-positive archetypes are represented by *E. coli* σ^S and *B. subtilis* σ^B, respectively (24, 31–33, 60). However, some bacterial groups completely lack homologs of the proteins involved in the above processes. In *Alphaproteobacteria*, an ECF-dependent regulatory cascade that combines 2CS- and ECF-dependent signaling mediates the general stress response by a mechanism that involves ECF mimicry and a partner-switching module (16, 18) (**Figure 2*e***).

The original ECF classification successfully identified all core functions of this general stress response, as mediated by group ECF15 σ factors: PhyR-like response regulators, NepR-like ASFs, EcfG-like ECFs, and at least one histidine kinase were usually encoded by neighboring genes in a single chromosomal location. Moreover, both the unique EcfG-like output domain of PhyR—which lacks the ability to bind DNA—and the specific signature sequence of NepR-like ASFs were identified, as was the target promoter signature (65, 66). All of these predictions were ultimately proven correct, as demonstrated by mechanistic studies from the Vorholt (18) and Crosson (16) groups.

PhyR-like response regulators are at the heart of these cascades. These proteins are unique in harboring their receiver domain at the C terminus and an EcfG-like output domain at the N terminus. The latter lacks crucial residues for DNA binding and indeed is not involved in transcription initiation. Instead, PhyR acts as a phosphorylation-dependent anti-ASF: In the absence of stress, PhyR closely folds in on itself, thereby masking the EcfG-like output domain. At the same time, the ASF NepR binds EcfG, thereby keeping it inactive (**Figure 2*e***). Upon stress induction, the kinase phosphorylates PhyR, causing a conformational change. PhyR releases the EcfG-like output domain from intramolecular inhibition, which, in turn, titrates NepR away from EcfG. As a result, this ECF becomes available for redirecting transcription initiation to the EcfG regulon (16, 18). To date, this represents the most complex of all ECF-dependent regulatory cascades and is also a beautiful example of how evolution has used the domains and modules of two unrelated signaling mechanisms, 2CSs and ECFs, to achieve a stimulus-integrating global response.

ECFs Controlled by Fused Regulatory Domains: The One-Component Systems of the ECF World

Of the 157 ECF groups defined in the current ECF classification, 43 lack discernable ASFs. Of these, 16 ECF groups harbor C-terminally fused, and 3 ECF groups harbor N-terminally fused, putative regulatory domains (11, 57). Since the identification of these groups by comparative

genomics, mechanistic studies of three of them have provided experimental evidence that these domains indeed play a regulatory role (**Figure 2*f***).

ECF41 is the most abundant and widely distributed ECF group and is found in 13 bacterial phyla. Its members are characterized by a C-terminal extension with a sensory SnoaL-like domain (57). Combined evidence from (*a*) a mutational analysis of ECF41 proteins from *R. sphaeroides* and *Bacillus licheniformis* (71), (*b*) a statistical analysis of the covariance between the extension and the ECF core (73), (*c*) the structure of *M. tuberculosis* SigJ (22), and (*d*) a structural simulation of the interaction dynamics between the ECF core and C-terminal extension (48) suggests a dual regulatory role of the SnoaL-like extension. In the absence of a stimulus, the C-terminal extension acts as an ASF-like domain by closely folding on the ECF core, thereby preventing its interaction with RNA polymerase. Upon perceiving a suitable trigger, conformational changes allow the ECF core to be recruited by the polymerase. But this process requires the core-proximal part of the extension in a manner that is not yet understood, since a complete deletion of the C-terminal extension results in a soluble but dysfunctional ECF41 core (71). Whether this positive regulatory role of the extension affects the ECF–polymerase or, rather, the holoenzyme–promoter interaction remains to be determined.

The physiological role of ECF41 proteins seems to be diverse. While *M. tuberculosis* SigJ has been implicated in resistance to hydrogen peroxide (36), RpoE10 from *Azospirillum brasilense* is indirectly involved in negatively regulating swimming motility and biogenesis of the lateral flagella (15). No phenotypes have been associated with the ECF41 proteins of *R. sphaeroides* and *B. licheniformis* (71). *Streptomyces tsukubaensis* SigG1, an ECF56 protein that also harbors a SnoaL-like extension, regulates morphogenesis and metal ion homeostasis during multicellular differentiation (53). In this ECF group, the C-terminal extension also interacts with the ECF core. Additionally, SigG1 is regulated by an ASF, RsiG (53), thereby demonstrating another combinatorial possibility of regulating ECF activity.

ECF42 is the second most abundant ECF group. It is distributed across 14 bacterial phyla and is especially enriched in Actinobacteria and Armatimonadetes (phylum *Armatimonadota*) genomes (73). ECF42 genes are typically located next to genes encoding DGPF proteins (also called YCII-related domain proteins) of unknown function, which are their primary targets. ECF42 proteins are characterized by 200–300-amino-acid-long C-terminal extensions harboring tetratricopeptide repeats, which are usually involved in protein–protein interactions. A direct coupling analysis predicted extensive interactions between the C terminus and the ECF core, which was confirmed by mutational studies (73). Gene deletion studies demonstrated that even minor truncations of the C termini lead to a complete loss of ECF activity (44). Together, these results suggest that the C termini of ECF42 proteins, while essential for σ factor activity, do not play an ASF-like role.

CorE and CorE2 of *M. xanthus* are the only members of group ECF238 (a merger of the original groups ECF24 and ECF44) that have been experimentally studied (21, 49, 52). They contain short (20-amino-acid), cysteine-rich C-terminal domains that are also found in other metal-binding proteins (57). While CorE is important for copper homeostasis, CorE2 mediates resistance against cadmium and zinc. In both cases, the metal ion specificity was determined by the C-terminal domain (21, 49, 52).

Taken together, these findings suggest that the C-terminal domains are ligand-binding domains essential for ECF activity that may or may not also harbor an ASF-like function.

σ Factor Phosphorylation: A Link Between ECFs and Ser/Thr Kinases

Transmembrane signal transduction usually involves separating a membrane-anchored sensor from a cytoplasmic regulator. This process necessitates a molecular means of communication in order to connect a certain stimulus to an appropriate cellular response. In bacterial 2CSs, such

communication is achieved by transient phosphotransfer reactions, while ECFs heavily rely on physical protein–protein interactions between a negative regulator and the σ factor (64). But the ECF classification indicates that σ factor phosphorylation may also be a possible way to activate ECFs. Currently, eight ECF groups are genomically linked to Ser/Thr kinases, while lacking obvious ASFs (11). A recent biochemical study demonstrated direct phosphorylation of the ECF43 σ factor EcfP from *Vibrio parahaemolyticus* by the membrane-anchored sensor Ser/Thr kinase PknT (**Figure 2*g***). EcfP is intrinsically inactive and can interact with the β′-subunit of RNA polymerase only upon direct phosphorylation. This leads to target gene expression, resulting in polymyxin resistance (38). This mechanism of ECF activation is therefore analogous to 2CS signaling.

Evolution of ECFs

ECFs developed presumably by reductive evolution from the more complex ancestral primary σ factors (23, 54, 58, 59). While ECFs are simple regulatory proteins with a minimalistic domain architecture, they have acquired increasingly complex accessory functions to control their activity (50, 58). While ASFs are the most common such additions, fused regulatory domains, phosphorylation-dependent mechanisms, and a combination of 2CS- and ECF-dependent signal transduction demonstrates the modular nature of bacterial signal transduction, which expresses itself at the level of protein domains rather than complete proteins. This evolutionary Lego found its most sophisticated manifestation in the Alphaproteobacterial general stress response, which functionally combines the receiver domain of a bacterial response regulator with an ECF-like output domain in PhyR-like proteins, in order to establish the partner-switching logic described above (65).

The ECF classification established that the vast majority of the ECF groups are phylum-specific, an indication that the ECF diversification is a relatively new development in evolution. Group ECF57, which is exclusively found in planctomycetes but is particularly enriched in only a few species, is an extreme case. *Gemmata obscuriglobus* alone encodes 62 ECF57 proteins that—while being highly homologous in their ECF core—show a remarkable variability in their long C-terminal extensions, which may contain few to many WD40-like β-propeller repeats and between zero and three transmembrane helices separating the N-terminal core from the C-terminal sensory domain (39). These proteins were presumably acquired once and then duplicated and permutated within the genome of this planctomycete to accommodate different physiological needs (56).

A comprehensive in silico study (56) on ECF evolution has demonstrated that single evolutionary events are the origin of alternative modes of ECF regulation that require specific partner proteins (such as ASFs or protein kinases), while multiple events resulted in the acquisition of regulatory extensions. Horizontal gene transfer of ECFs was also documented in the distribution of group ECF20 in gram-positive streptomycetes, which were originally acquired from the gram-negative *Sinorhizobium* proteobacterium (56).

THE FUTURE: CURRENT LIMITATIONS, CHALLENGES, AND OPPORTUNITIES OF ECF RESEARCH

ECFs were recognized 30 years ago as a novel group of σ^{70} proteins: The first multiple-sequence alignment of three members of the ECF subfamily by Michael Lonetto dates back to May 19, 1993 (46). Since the seminal study on defining the ECF subfamily (45), approximately 650 articles on the topic have been published and can be extracted from PubMed (**https://pubmed.ncbi.nlm.nih.gov**) with the search string "extracytoplasmic function sigma factors" (as of June 2023). The number of published papers on ECFs has steadily increased from 1994 to 2005. Since then, approximately 25–30 papers on this subject have been published each year.

While the ECF classification efforts did not increase the number of ECF studies, they helped identify novel mechanisms of controlling ECF activity, as outlined above. Most importantly, comparative genomics paved the way toward the future of ECF research by providing comprehensive and reliable in silico support, both for mechanistic studies on ECF regulation and for unraveling their physiological role.

"Way Out Yonder Where the Crawdads Sing": The Unexplored World of ECF Signaling

The goal of this review has been to summarize the diversity of ECF-dependent signaling from the comparative genomics perspective of the ECF classification. This in silico effort enables the identification of novel regulatory mechanisms for ECF groups that have not been studied experimentally. The reliability and predictive power of this classification have been demonstrated for the EcfG-dependent general stress response (ECF15), ECFs with C-terminal extensions (ECF41, ECF42, ECF238), and phosphorylation-dependent activation of ECFs (ECF43), as described above. But the latest ECF classification provides numerous additional examples, as described in the ECF Hub (see Related Resources). Hopefully, these examples will inspire many future studies on altogether unusual modes of ECF regulation. Some promising candidates are described next.

The current classification defines 157 ECF groups, of which 114 contain a putative ASF (11). The vast majority of these harbor a single membrane-spanning region, thereby presumably following the classical blueprint of regulated proteolysis, as established for *E. coli* σ^E-RseA or *B. subtilis* σ^W-RsiW (1, 2, 27, 63). But even for ECFs associated with archetypical ASFs, initial studies indicate a remarkable regulatory diversity. Members of group ECF102 are encoded in an operon together with an ASF and a mechanosensitive ion channel. The only member of this group studied so far, σ^X of *P. aeruginosa*, has been implicated in mechanosensing, involving the ASF CfrX, the mechanosensitive channel CmpX, and the outer membrane porin OprF (12). Additional, accessory proteins that control ECF activity together with the ASF can be found in group ECF31, in which two proteins are encoded in an operon together with the ECF gene. Both seem to be involved in controlling ECF activity, as demonstrated for *B. subtilis* σ^Y, which requires two ASFs, YxlC and YxlD, for regulation (75). Ten ECF groups are linked to soluble ASFs, as exemplified by the paradigms *R. sphaeroides* σ^E-ChrR and *S. coelicolor* σ^R-RsrA (14, 55). In addition, 12 ECF groups are associated with ASFs harboring two or three transmembrane helices, while 8 ECF groups are linked to ASFs with four or six such regions (11). While an untapped regulatory diversity may lay hidden in all of these ECF groups, the 43 ECF groups lacking obvious ASFs are even more thrilling, since they necessitate alternative regulatory mechanisms.

Fused regulatory extensions represent the second most common mode of ECF control and can be found in 16 ECF groups, including some of the largest, phylogenetically most widespread groups, such as ECF41 and ECF42, described above (57). Many contain conserved domains in their extensions, such as tetratricopeptide repeats in group ECF42 or SnoaL-like extensions in groups ECF41, ECF56, ECF294, and ECF295 (11). These examples indicate that the regulatory mechanism most likely involves extensive intramolecular interactions between the extension and the ECF core region to regulate ECF activity (73). Both ASF-like functions and positive regulatory roles, which most likely act through direct interactions with the RNA polymerase, have been observed for ECF41 and ECF42 (73).

Some groups containing C-terminal extensions harbor membrane-spanning domains between the N-terminal ECF core and the sensory unit (e.g., ECF264). How these are activated is purely speculative, but release of the N-terminal ECF domains into the cytoplasm after proteolytic cleavage from the membrane-anchored/extracellular C terminus is a very attractive option.

Three ECF groups from the planctomycetes harbor large N-terminal regulatory domains. Their regulatory mechanism has yet to be unraveled. In addition, some ECF groups harbor short N-terminal extensions that have indeed been associated with a regulatory role. The ECF121 member BldN from *S. coelicolor* is proteolytically processed at its N terminus to yield its mature form, which is then controlled by an ASF (3). Some members of group ECF36 lack a discernable ASF, but the N-terminal extension of one of its members, *M. tuberculosis* SigC, was proposed to inhibit DNA contact in the uninduced state (68). In fact, one of the paradigmatic ECFs mentioned above, the group ECF12 member σ^R from *S. coelicolor*, is produced as an instable isoform from an earlier start codon upon exposure to thiol oxidants, which adds a negative regulatory feedback loop by making this isoform susceptible to σ^R-mediated, ClpP1/P2-dependent proteolysis (41).

Phosphorylation-dependent control of ECFs is another widespread mechanism: The ECF classification lists nine ECF groups that are genomically linked to Ser/Thr kinases, six of which are not associated with any putative ASF. This finding indicates that at least those six groups—if not all nine—might also be subject to direct ECF phosphorylation, as described for ECF43 (38). Beyond the ECF15-mediated general stress response, 11 additional ECF groups are genomically linked to 2CSs. All but two of them harbor putative ASF genes next to those encoding the ECF, indicative of additional regulatory links between 2CS- and ECF/ASF-dependent signal transduction. The ECF282 member SigP from *Porphyromonas gingivalis* is associated with and stabilized by the response regulator PorX of the 2CS PorXY (40).

This leaves 16 groups not linked to any of the mechanisms described above. Some ECFs might be subject to transcriptional regulation, such as the ECF114 member SigH from *P. gingivalis* (74). Others, such as ECF203, are linked to TetR-like transcriptional regulators, again suggesting transcriptional control. The ECF282 member σ^{AntA} is an orphan ECF that is transcriptionally regulated but also subject to ClpXP proteolysis, which adds another layer of control (62). Some of the remaining groups show genomic context conservation to genes encoding proteins with helix-turn-helix motifs (ECF130, ECF201), proteins with a 4Fe-4S cluster (ECF54), or Asp23 proteins (ECF286, ECF292). But whether these are regulatory functions or target genes is hard to determine without experimental studies.

"Here Be Dragons": The Phylogenetic Terra Incognita of ECF-Dependent Signaling in Underrepresented Bacterial Phyla

Even the most current ECF (re)classification effort, as represented in the ECF Hub (see Related Resources), is still far from being representative of the full phylogenetic diversity of the bacterial world, due to an intrinsic bias of the microbial genomic sequence space toward highly represented phylogenetic groups: 93% of all ECFs classified to date are derived from only four bacterial phyla (*Proteobacteria*, Actinobacteria, *Firmicutes*, *Bacteroidetes*), whereas only 7% cover the remaining 127 phyla (**Figure 3*a***). Advances in metagenomic and single-cell sequencing techniques have greatly expanded the diversity of the NCBI microbial genome sequence database, which prompted a closer examination of these underrepresented phyla, including sequences from metagenomic data, in order to complete the comparative genomics perspective on ECF diversity. The results of this so-far-unpublished study are briefly summarized below. For more details, including the methods used, see the **Supplemental Material**.

Supplemental Material >

Out of 127,846 genomes from all 132 bacterial phyla, 8,859 genomes belong to the 127 underrepresented phyla. Restricting our analysis to genomes that are at least 90% complete yielded 4,212 genomes from 117 phyla, of which 2,987 (from 109 phyla) encoded ECFs. No ECFs were found in eight novel ("*Candidatus*") phyla, while the average number of ECFs per genome was below one in another 11 phyla (e.g., *Aquificota*, *Tenericutes*). In contrast, some of the underrepresented

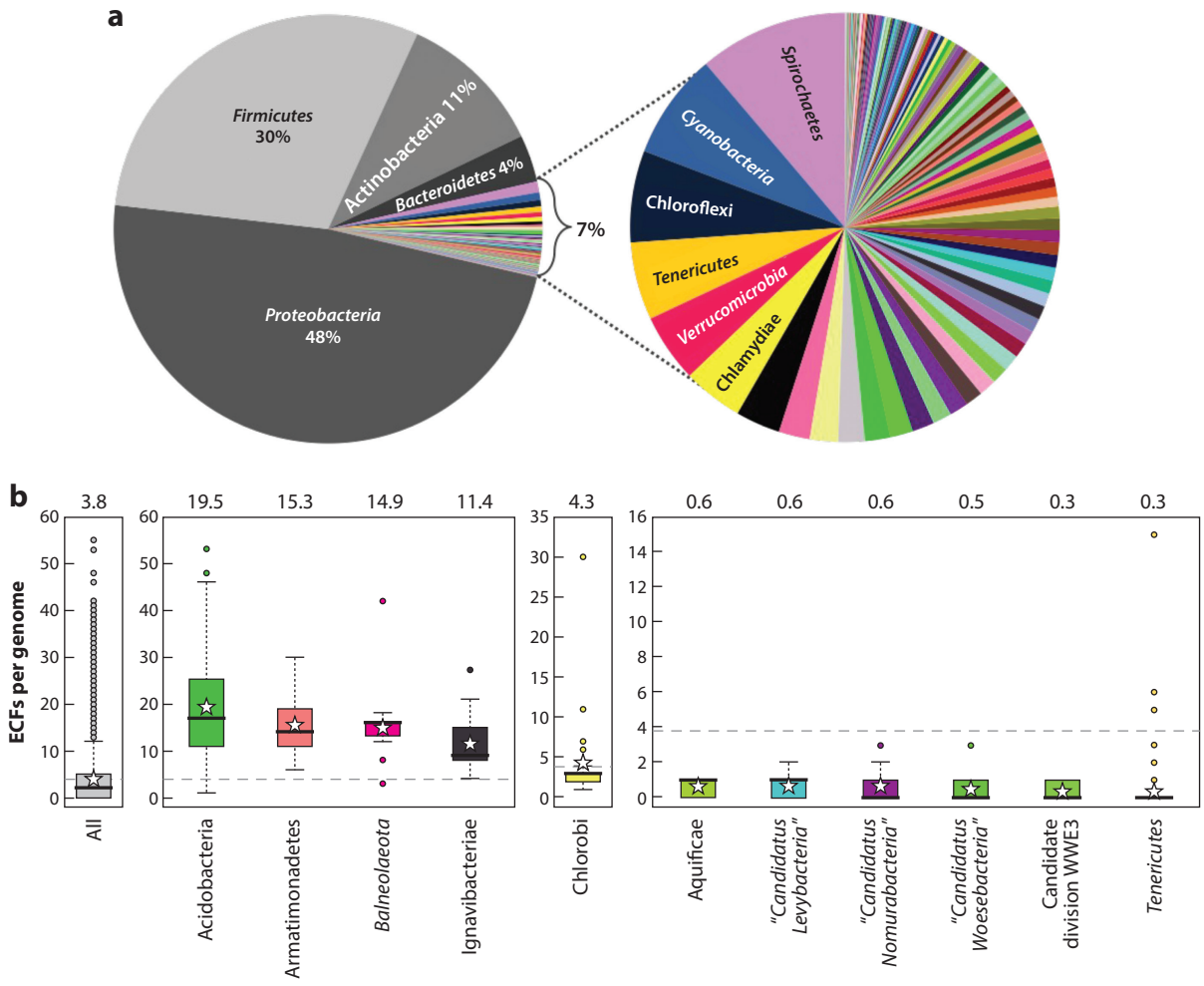

Figure 3

(*a*) Genome sequence bias in microbial genome databases. Only four predominant bacterial phyla account for 93% of all genome sequences, while the remaining 127 phyla (*highlighted in color*) are represented by only 7% of the genome sequences. (*b*) The box plots illustrate the number of ECFs found per genome in a selection of different phyla; the numbers vary greatly depending on the phylum. In each box plot, each white star represents the mean number of ECFs within the given phylum, and the corresponding value is depicted on top of the graph. The first box plot shows the number of ECFs in each genome across all phyla with a mean value of 3.8 ECFs per genome, which is shown as a gray dotted line in the three other graphs. The second box plot shows four phyla with a mean number of ECFs per genome well above the average of 3.8 ECFs. The third box plot highlights the phylum *Chlorobiota*, which represents the average found over all phyla. Abbreviation: ECF, extracytoplasmic function σ factor.

phyla, such as *Acidobacteriota*, *Armatimonadota*, *Balneolaeota*, and *Ignavibacteriota*, are particularly ECF-rich (**Figure 3*b***). A total of 16,181 ECFs were retrieved and 15,346 sequences were associated to 1 of 841 different clusters. Seventy-five percent of these clusters contained fewer than 10 sequences, but together they accounted for only 15% of all clustered ECFs. Twenty-two percent of all clusters contained 10–100 ECF sequences and accounted for 31% of all ECFs, while 3% of the clusters contained more than 100 sequences but accounted for 54% of all ECFs in the data set.

Supplemental Material >

Scanning this ECF data set with the group-specific hidden Markov models from the current reclassification (11) allowed 71% of all ECFs from underrepresented phyla to be classified and assigned to 1 of 85 ECF groups. Of the 29% unclassified ECFs, all clusters with more than 10 sequences were subjected to in-depth analyses, following the procedure published elsewhere (37, 66, 72). These analyses resulted in the definition of 43 novel ECF groups, ECF140–ECF182, that go beyond the recent reclassification effort. Most of these groups were found exclusively in underrepresented phyla, and many were restricted to one phylum only. For a detailed description of these groups, see the **Supplemental Material**.

A total of 15 groups are associated with ASFs, while C-terminal extensions can be found in 4 groups (ECF140, ECF146, ECF175, and ECF182), with ECF146 members containing six putative transmembrane helices between the ECF core and the extension. ECF160/ECF173 σ factors are associated with Clp proteases, while ECF147 might be controlled by phosphorylation. A clear group-specific promoter signature has been predicted for at least 10 ECF groups (see the **Supplemental Material**), thereby enabling regulon predictions.

With this final classification effort, and the information provided by the ECF Hub, the diversity of ECF-dependent signal transduction and gene regulation—as classified in 200 distinct ECF groups (**Table 1**)—has now been comprehensively documented and is available to support future in-depth mechanistic studies of individual members from any of these groups.

SUMMARY POINTS

1. ECFs are the simplest and most diverse members of the σ^{70} protein family.
2. ECFs represent the third most abundant—and by far the most diverse—mechanism of bacterial signal transduction.
3. ECFs can be controlled by membrane-anchored or soluble ASFs, regulatory extensions, σ factor phosphorylation, and σ factor mimicry as well as at the transcriptional level.
4. The current ECF classification defines 200 ECF groups, based on their sequence similarity, ASFs, genomic context conservation, and target promoter motif.
5. The ECF classification provides a comprehensive and reliable resource for predicting (novel) mechanisms of ECF regulation and determining their physiological role.

FUTURE ISSUES

1. Our understanding of the diversity of ECF regulation is far from complete, since the majority of ECF groups, including many ECFs predicted to be controlled by altogether novel regulatory mechanisms, have not yet been experimentally addressed.
2. The inherent genome sequence bias still hampers the comprehensiveness of the ECF classification for the vast majority of underrepresented bacterial phyla.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

With this overview, I close a chapter and leave ECF research behind. What we could contribute, we did. With the ECF classification, we hope to inspire many mechanistic studies on novel ECF groups for years to come. I thank my mentor, John D. Helmann (Cornell University), for putting me on the ECF train 20 years ago. Celebrating 25 years of ECF research with him in 2018–2019 was a fulfilling way of concluding this long collaboration. Moreover, I thank all past members of the Mascher group involved in ECF research (in order of appearance): Tina Wecke, Anna Nagy-Stáron, Georg Fritz, Franziska Dürr (who contributed the ECF classification of underrepresented phyla provided in this review), Xiaoluo Huang, Daniela Pinto, Dayane Araújo, and Qiang Liu. This review is dedicated to Kya, who inspired me and contributed an important aspect of this article. Since its focus is on ECF diversity rather than on the details of regulatory mechanisms, it relies heavily on overview articles. I therefore thank and apologize to all colleagues whose groundbreaking experimental studies on unraveling the regulatory mechanisms of all the different ECFs briefly touched on herein are not cited in the context of this article. Over the past 15 years, research on ECFs in my group has been supported by grants from the Deutsche Forschungsgemeinschaft, the Federal Ministry of Education and Research (in the context of ERAsynbio), the Marie Curie program of the European Union (to Daniela Pinto), the Graduate Program of TU Dresden (to Franziska Dürr), and the China Scholarship Council (to Xiaoluo Huang and Qiang Liu).

LITERATURE CITED

1. Ades SE. 2008. Regulation by destruction: design of the σ^E envelope stress response. *Curr. Opin. Microbiol.* 11:535–40
2. Barchinger SE, Ades SE. 2013. Regulated proteolysis: control of the *Escherichia coli* σ^E-dependent cell envelope stress response. *Subcell. Biochem.* 66:129–60
3. Bibb MJ, Buttner MJ. 2003. The *Streptomyces coelicolor* developmental transcription factor σ^{BldN} is synthesized as a proprotein. *J. Bacteriol.* 185:2338–45
4. Bordi C, Butcher BG, Shi Q, Hachmann AB, Peters JE, Helmann JD. 2008. In vitro mutagenesis of *Bacillus subtilis* by using a modified Tn7 transposon with an outward-facing inducible promoter. *Appl. Environ. Microbiol.* 74:3419–25
5. Braun V, Mahren S, Sauter A. 2006. Gene regulation by transmembrane signaling. *Biometals* 19:103–13
6. Brooks BE, Buchanan SK. 2008. Signaling mechanisms for activation of extracytoplasmic function (ECF) sigma factors. *Biochim. Biophys. Acta Biomembr.* 1778:1930–45
7. Browning DF, Busby SJ. 2016. Local and global regulation of transcription initiation in bacteria. *Nat. Rev. Microbiol.* 14:638–50
8. Buchner S, Schlundt A, Lassak J, Sattler M, Jung K. 2015. Structural and functional analysis of the signal-transducing linker in the pH-responsive one-component system CadC of *Escherichia coli*. *J. Mol. Biol.* 427:2548–61
9. Butcher BG, Mascher T, Helmann JD. 2008. Environmental sensing and the role of extracytoplasmic function (ECF) sigma factors. In *Bacterial Physiology: A Molecular Approach*, ed. WM El-Sharoud, pp. 233–61. Berlin/Heidelberg: Springer
10. Campbell EA, Greenwell R, Anthony JR, Wang S, Lim L, et al. 2007. A conserved structural module regulates transcriptional responses to diverse stress signals in bacteria. *Mol. Cell* 27:793–805
11. Casas-Pastor D, Müller RR, Jaenicke S, Brinkrolf K, Becker A, et al. 2021. Expansion and re-classification of the extracytoplasmic function (ECF) sigma factor family. *Nucleic Acids Res.* 49:986–1005
12. Chevalier S, Bouffartigues E, Bazire A, Tahrioui A, Duchesne R, et al. 2019. Extracytoplasmic function sigma factors in *Pseudomonas aeruginosa*. *Biochim. Biophys. Acta Gene Regul. Mech.* 1862:706–21
13. Darnell RL, Nakatani Y, Knottenbelt MK, Gebhard S, Cook GM. 2019. Functional characterization of BcrR: a one-component transmembrane signal transduction system for bacitracin resistance. *Microbiology* 165:475–87
14. Donohue TJ. 2019. Shedding light on a group IV (ECF11) alternative σ factor. *Mol. Microbiol.* 112:374–84

15. Dubey AP, Pandey P, Singh VS, Mishra MN, Singh S, et al. 2020. An ECF41 family σ factor controls motility and biogenesis of lateral flagella in *Azospirillum brasilense* Sp245. *J. Bacteriol.* 202:e00231
16. Fiebig A, Herrou J, Willett J, Crosson S. 2015. General stress signaling in the Alphaproteobacteria. *Annu. Rev. Genet.* 49:603–25
17. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, et al. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269:496–512
18. Francez-Charlot A, Kaczmarczyk A, Fischer HM, Vorholt JA. 2015. The general stress response in Alphaproteobacteria. *Trends Microbiol.* 23:164–71
19. Fraser CM, Gocayne JD, White O, Adams MD, Clayton RA, et al. 1995. The minimal gene complement of *Mycoplasma genitalium*. *Science* 270:397–403
20. Gao R, Stock AM. 2009. Biological insights from structures of two-component proteins. *Annu. Rev. Microbiol.* 63:133–54
21. Gómez-Santos N, Pérez J, Sánchez-Sutil MC, Moraleda-Muñoz A, Muñoz-Dorado J. 2011. CorE from *Myxococcus xanthus* is a copper-dependent RNA polymerase sigma factor. *PLOS Genet.* 7:e1002106
22. Goutam K, Gupta AK, Gopal B. 2017. The fused SnoaL_2 domain in the *Mycobacterium tuberculosis* sigma factor σ^J modulates promoter recognition. *Nucleic Acids Res.* 45:9760–72
23. Gruber TM, Gross CA. 2003. Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu. Rev. Microbiol.* 57:441–66
24. Hecker M, Pane-Farre J, Völker U. 2007. σ^B-dependent general stress response in *Bacillus subtilis* and related gram-positive bacteria. *Annu. Rev. Microbiol.* 61:215–36
25. Heinrich J, Wiegert T. 2009. Regulated intramembrane proteolysis in the control of extracytoplasmic function sigma factors. *Res. Microbiol.* 160:696–703
26. Helmann JD. 2002. The extracytoplasmic function (ECF) sigma factors. *Adv. Microb. Physiol.* 46:47–110
27. Helmann JD. 2006. Deciphering a complex genetic regulatory network: the *Bacillus subtilis* σ^W protein and intrinsic resistance to antimicrobial compounds. *Sci. Prog.* 89:243–66
28. Helmann JD. 2016. *Bacillus subtilis* extracytoplasmic function (ECF) sigma factors and defense of the cell envelope. *Curr. Opin. Microbiol.* 30:122–32
29. Helmann JD. 2019. Where to begin? Sigma factors and the selectivity of transcription initiation in bacteria. *Mol. Microbiol.* 112:335–47
30. Helmann JD, Chamberlin MJ. 1988. Structure and function of bacterial sigma factors. *Annu. Rev. Biochem.* 57:839–72
31. Hengge R. 2008. The two-component network and the general stress sigma factor RpoS (σ^S) in *Escherichia coli*. *Adv. Exp. Med. Biol.* 631:40–53
32. Hengge R. 2009. Proteolysis of σ^S (RpoS) and the general stress response in *Escherichia coli*. *Res. Microbiol.* 160:667–76
33. Hengge R. 2011. The general stress response in gram-negative bacteria. In *Bacterial Stress Responses*, ed. G Storz, R Hengge, pp. 251–90. Washington, DC: ASM
34. Ho TD, Ellermeier CD. 2012. Extracytoplasmic function sigma factor activation. *Curr. Opin. Microbiol.* 15:182–88
35. Ho TD, Ellermeier CD. 2019. Activation of the extracytoplasmic function sigma factor σ^V by lysozyme. *Mol. Microbiol.* 112:410–19
36. Hu Y, Kendall S, Stoker NG, Coates AR. 2004. The *Mycobacterium tuberculosis sigJ* gene controls sensitivity of the bacterium to hydrogen peroxide. *FEMS Microbiol. Lett.* 237:415–23
37. Huang XL, Pinto D, Fritz G, Mascher T. 2015. Environmental sensing in Actinobacteria: a comprehensive survey on the signaling capacity of this phylum. *J. Bacteriol.* 197:2517–35
38. Iyer SC, Casas-Pastor D, Kraus D, Mann P, Schirner K, et al. 2020. Transcriptional regulation by σ factor phosphorylation in bacteria. *Nat. Microbiol.* 5:395–406
39. Jogler C, Waldmann J, Huang X, Jogler M, Glöckner FO, et al. 2012. Identification of proteins likely to be involved in morphogenesis, cell division, and signal transduction in planctomycetes by comparative genomics. *J. Bacteriol.* 194:6419–30
40. Kadowaki T, Yukitake H, Naito M, Sato K, Kikuchi Y, et al. 2016. A two-component system regulates gene expression of the type IX secretion component proteins via an ECF σ factor. *Sci. Rep.* 6:23288

41. Kim MS, Hahn MY, Cho Y, Cho SN, Roe JH. 2009. Positive and negative feedback regulatory loops of thiol-oxidative stress response mediated by an unstable isoform of σ^R in actinomycetes. *Mol. Microbiol.* 73:815–25
42. Li LT, Fang CL, Zhuang NN, Wang TT, Zhang Y. 2019. Structural basis for transcription initiation by bacterial ECF σ factors. *Nat. Commun.* 10:1153
43. Lin W, Mandal S, Degen D, Cho MS, Feng Y, et al. 2019. Structural basis of ECF-σ-factor-dependent transcription initiation. *Nat. Commun.* 10:710
44. Liu Q, Pinto D, Mascher T. 2018. Characterization of the widely distributed novel ECF42 group of extracytoplasmic function σ factors in *Streptomyces venezuelae*. *J. Bacteriol.* 200:e00437
45. Lonetto MA, Brown KL, Rudd KE, Buttner MJ. 1994. Analysis of the *Streptomyces coelicolor sigE* gene reveals the existence of a subfamily of eubacterial RNA polymerase sigma factors involved in the regulation of extracytoplasmic functions. *PNAS* 91:7573–77
46. Lonetto MA, Donohue TJ, Gross CA, Buttner MJ. 2019. Discovery of the extracytoplasmic function σ factors. *Mol. Microbiol.* 112:348–55
47. Mahren S, Schnell H, Braun V. 2005. Occurrence and regulation of the ferric citrate transport system in *Escherichia coli* B, *Klebsiella pneumoniae*, *Enterobacter aerogenes*, and *Photorhabdus luminescens*. *Arch. Microbiol.* 184:175–86
48. Mallick Gupta A, Mandal S. 2020. The C-terminal domain of *M. tuberculosis* ECF σ factor I (SigI) interferes in SigI-RNAP interaction. *J. Mol. Model.* 26:77
49. Marcos-Torres FJ, Pérez J, Gomez-Santos N, Moraleda-Muñoz A, Muñoz-Dorado J. 2016. In depth analysis of the mechanism of action of metal-dependent σ factors: characterization of CorE2 from *Myxococcus xanthus*. *Nucleic Acids Res.* 44:5571–84
50. Mascher T. 2013. Signaling diversity and evolution of extracytoplasmic function (ECF) σ factors. *Curr. Opin. Microbiol.* 16:148–55
51. Missiakas D, Raina S. 1998. The extracytoplasmic function sigma factors: role and regulation. *Mol. Microbiol.* 28:1059–66
52. Moraleda-Muñoz A, Marcos-Torres FJ, Pérez J, Muñoz-Dorado J. 2019. Metal-responsive RNA polymerase extracytoplasmic function (ECF) sigma factors. *Mol. Microbiol.* 112:385–98
53. Oliveira R, Bush MJ, Pires S, Chandra G, Casas-Pastor D, et al. 2020. The novel ECF56 SigG1-RsfG system modulates morphological differentiation and metal-ion homeostasis in *Streptomyces tsukubaensis*. *Sci. Rep.* 10:21728
54. Paget MS, Helmann JD. 2003. The σ^{70} family of sigma factors. *Genome Biol.* 4:203
55. Park JH, Lee JH, Roe JH. 2019. SigR, a hub of multilayered regulation of redox and antibiotic stress responses. *Mol. Microbiol.* 112:420–31
56. Pinto D, da Fonseca RR. 2020. Evolution of the extracytoplasmic function σ factor protein family. *NAR Genom. Bioinform.* 2:lqz026
57. Pinto D, Liu Q, Mascher T. 2019. ECF σ factors with regulatory extensions: the one-component systems of the σ universe. *Mol. Microbiol.* 112:399–409
58. Pinto D, Mascher T. 2016. (Actino)bacterial "intelligence": using comparative genomics to unravel the information processing capacities of microbes. *Curr. Genet.* 62:487–98
59. Pinto D, Mascher T. 2016. The ECF classification: a phylogenetic reflection of the regulatory diversity in the extracytoplasmic function σ factor protein family. In *Stress and Environmental Regulation of Gene Expression and Adaptation in Bacteria*, ed. FJ De Bruijn, Vol. 1, pp. 64–96. Hoboken, NJ: Wiley Blackwell
60. Price CW. 2011. The general stress response in *Bacillus subtilis* and related gram-positive bacteria. In *Bacterial Stress Responses*, ed. G Storz, R Hengge, pp. 301–18. Washington, DC: ASM
61. Schneider JS, Glickman MS. 2013. Function of site-2 proteases in bacteria and bacterial pathogens. *Biochim. Biophys. Acta Biomembr.* 1828:2808–14
62. Seipke RF, Patrick E, Hutchings MI. 2014. Regulation of antimycin biosynthesis by the orphan ECF RNA polymerase σ factor σ^{AntA}. *PeerJ* 2:e253
63. Sineva E, Savkina M, Ades SE. 2017. Themes and variations in gene regulation by extracytoplasmic function (ECF) σ factors. *Curr. Opin. Microbiol.* 36:128–37
64. Staroń A, Mascher T. 2010. Extracytoplasmic function σ factors come of age. *Microbe* 5:164–70

65. Staroń A, Mascher T. 2010. General stress response in α-proteobacteria: PhyR and beyond. *Mol. Microbiol.* 78:271–77
66. Staroń A, Sofia HJ, Dietrich S, Ulrich LE, Liesegang H, Mascher T. 2009. The third pillar of bacterial signal transduction: classification of the extracytoplasmic function (ECF) σ factor protein family. *Mol. Microbiol.* 74:557–81
67. Stock AM, Robinson VL, Goudreau PN. 2000. Two-component signal transduction. *Annu. Rev. Biochem.* 69:183–215
68. Thakur KG, Joshi AM, Gopal B. 2007. Structural and biophysical studies on two promoter recognition domains of the extra-cytoplasmic function σ factor σ^C from *Mycobacterium tuberculosis*. *J. Biol. Chem.* 282:4711–18
69. Tran NT, Huang X, Hong HJ, Bush MJ, Chandra G, et al. 2019. Defining the regulon of genes controlled by σ^E, a key regulator of the cell envelope stress response in *Streptomyces coelicolor*. *Mol. Microbiol.* 112:461–81
70. Ulrich LE, Koonin EV, Zhulin IB. 2005. One-component systems dominate signal transduction in prokaryotes. *Trends Microbiol.* 13:52–56
71. Wecke T, Halang P, Staron A, Dufour YS, Donohue TJ, Mascher T. 2012. Extracytoplasmic function σ factors of the widely distributed group ECF41 contain a fused regulatory domain. *MicrobiologyOpen* 1:194–213
72. Wiegand S, Jogler M, Boedeker C, Pinto D, Vollmers J, et al. 2020. Cultivation and functional characterization of 79 planctomycetes uncovers their unique biology. *Nat. Microbiol.* 5:126–40
73. Wu H, Liu Q, Casas-Pastor D, Durr F, Mascher T, Fritz G. 2019. The role of C-terminal extensions in controlling ECF σ factor activity in the widely conserved groups ECF41 and ECF42. *Mol. Microbiol.* 112:498–514
74. Yanamandra SS, Sarrafee SS, Anaya-Bergman C, Jones K, Lewis JP. 2012. Role of the *Porphyromonas gingivalis* extracytoplasmic function σ factor, SigH. *Mol. Oral Microbiol.* 27:202–19
75. Yoshimura M, Asai K, Sadaie Y, Yoshikawa H. 2004. Interaction of *Bacillus subtilis* extracytoplasmic function (ECF) sigma factors with the N-terminal regions of their potential anti-sigma factors. *Microbiology* 150:591–99

RELATED RESOURCES

ECF Hub (**https://www.uni-giessen.de/de/fbz/fb08/Inst/bioinformatik/software/ECF%20Hub**). The ECF Hub is a comprehensive resource for everything ECF. It provides access to the current ECF classification, an ECF literature database, detailed descriptions of individual ECF groups, and sequence analysis tools (11).