

REVIEW

# Two-dimensional gel electrophoresis in bacterial proteomics

Shirly O. T. Curreem<sup>1\*</sup>, Rory M. Watt<sup>2\*</sup>, Susanna K. P. Lau<sup>1,3,4,5</sup>✉, Patrick C. Y. Woo<sup>1,3,4,5</sup>✉

<sup>1</sup> Department of Microbiology, The University of Hong Kong, Queen Mary Hospital, Hong Kong, China

<sup>2</sup> Oral Biosciences, Faculty of Dentistry, The University of Hong Kong, Queen Mary Hospital, Hong Kong, China

<sup>3</sup> State Key Laboratory of Emerging Infectious Diseases, Department of Microbiology, The University of Hong Kong, Queen Mary Hospital, Hong Kong, China

<sup>4</sup> Research Centre of Infection and Immunology, The University of Hong Kong, Queen Mary Hospital, Hong Kong, China

<sup>5</sup> Carol Yu Centre of Infection, The University of Hong Kong, Hong Kong, China

✉ Correspondence: pcywoo@hkucc.hku.hk (P. C. Y. Woo), skplau@hkucc.hku.hk (S. K. P. Lau)

Received February 4, 2012 Accepted February 22, 2012

## ABSTRACT

Two-dimensional gel electrophoresis (2-DE) is a gel-based technique widely used for analyzing the protein composition of biological samples. It is capable of resolving complex mixtures containing more than a thousand protein components into individual protein spots through the coupling of two orthogonal biophysical separation techniques: isoelectric focusing (first dimension) and polyacrylamide gel electrophoresis (second dimension). 2-DE is ideally suited for analyzing the entire expressed protein complement of a bacterial cell: its proteome. Its relative simplicity and good reproducibility have led to 2-DE being widely used for exploring proteomics within a wide range of environmental and medically-relevant bacteria. Here we give a broad overview of the basic principles and historical development of gel-based proteomics, and how this powerful approach can be applied for studying bacterial biology and physiology. We highlight specific 2-DE applications that can be used to analyze when, where and how much proteins are expressed. The links between proteomics, genomics and mass spectrometry are discussed. We explore how proteomics involving tandem mass spectrometry can be used to analyze (post-translational) protein modifications or to identify proteins of unknown origin by *de novo* peptide sequencing. The use of proteome fractionation techniques and non-gel-based proteomic approaches are

also discussed. We highlight how the analysis of proteins secreted by bacterial cells (secretomes or exoproteomes) can be used to study infection processes or the immune response. This review is aimed at non-specialists who wish to gain a concise, comprehensive and contemporary overview of the nature and applications of bacterial proteomics.

**KEYWORDS** two-dimensional gel electrophoresis, bacteria, proteomics

## INTRODUCTION

The genomic era began when the first bacterial genomic sequence, that of *Haemophilus influenzae* Rd, was completed in 1995 (Fleischmann et al., 1995). With ongoing advancements in sequencing technologies, the ability to sequence entire genomes has become progressively easier, faster and cheaper to achieve. This has resulted in an exponential growth in the number of species whose genomes have been sequenced, which greatly advanced our understanding of the natural world. Furthermore, genome-based approaches have revolutionized our whole approach to study organismal biology, especially that of bacteria. Sequencing the genome sequence of a bacterial species or an isolate enables us to gain insight into its biochemistry and cellular physiology, and an understanding of how it may adapt to its specific ecological niches. It also gives us the ability to study its evolutionary history and genetic relationships to other

\*These authors contributed equally to the work.

organisms. Genomic information has led to massive advancements in drug and vaccine discovery, molecular epidemiology, diagnostic bacteriology and knowledge of antimicrobial resistance mechanisms (Rappuoli, 2001; Miesel et al., 2003; Mora et al., 2006; Parkhill and Wren, 2011)

The genome is a static phase that contains a set number of genes that encode proteins or functional RNA molecules, e.g. ribosomal RNA (rRNA), transfer RNA (tRNA) or non-coding RNA (ncRNA). Sequencing an organism's DNA complement reveals its genetic makeup; however, it is proteins that mediate the vast majority of cellular processes. It is often inaccurate to depend solely on genomic information to infer the functions of proteins, or to rely purely on transcriptomic data (i.e. the production of messenger RNA) to quantify protein expression levels. Cellular protein expression patterns are highly dynamic, in terms of how they are synthesized, degraded, modified, or distributed within or without the cell. It is more accurate and informative, yet often more technically-demanding, to study proteins directly.

The term 'proteome' was first coined by Marc Wilkins in 1994, to describe the entire set of proteins encoded by the genome of an organism, cell or tissue at a given time (Wilkins et al., 1996). The study of proteomes is termed 'proteomics,' and may involve a variety of different aspects related to the study of proteins at a global or organismal level. This includes the analysis of protein expression patterns, biological functions, macromolecular protein structures, spatiotemporal intracellular distributions, stabilities and turnover rates, post-translational protein modifications as well as protein-protein interactions. 'Modern' approaches towards the large-scale study of proteins began in the mid-1970s, when a technique known as two-dimensional electrophoresis (2-DE) was first developed (Klose, 1975; O'Farrell, 1975). 2-DE combines two different separation methods: isoelectric focusing (IEF) and 'conventional' sodium-dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). This approach enables complex mixtures containing many hundreds to thousands of different proteins to be spatially-resolved in two-dimensional space using a gel-based matrix. Individual protein components may be subsequently visualized as 'spots' on the gels, which have diameters of a few millimeters (gels used are generally Y cm × Z cm). Developments in 2-DE over the last thirty years or so have transformed protein-based research from a hypothesis-driven to a discovery-driven science. Boosted greatly by the availability of complete genome sequences, comprehensive protein databases and powerful computational tools, large-scale protein analysis has become increasingly well-used for deciphering protein function and identifying protein networks and regulatory mechanisms. Nowadays, proteomic investigations are generally conducted using three general approaches: (1) gel-based techniques, (2) non-gel-based techniques, and (3) computational (predictive) approaches. The former two approaches are based on the separation of complex protein

mixtures using gel or non gel-based matrices, whereas predictive proteomics are performed entirely *in silico*. In this review, we give a historical overview of the development of 2-DE methods. We summarize the major 2-DE applications used for studying bacterial proteomes, and highlight their respective advantages and disadvantages. We also briefly discuss several non gel-based approaches that may be used to complement gel-based techniques.

## HISTORICAL OVERVIEW OF 2-DE

In the 1950s, it was well recognized that traditional one-dimensional gel-based separations were incapable of resolving complex protein mixtures into individual component species (Klose, 2009). Independently developed by O'Farrell, Klose and Scheele in 1975, 2-DE is a powerful gel-based technique that enables the separation of complex protein mixtures into individual protein 'spots,' based on their isoelectric point (pI) and molecular weight (MW) (Klose, 1975; O'Farrell, 1975; Scheele, 1975). In one of these pioneering studies, O'Farrell successfully separated *Escherichia coli* cell extracts into thousands of protein components, representing the first example of 'bacterial proteomics' using the 2-DE approach (O'Farrell, 1975). 2-DE is a gel-based approach widely used for analyzing the protein composition of a given biological sample. As suggested by its name, this technique involves the use of two orthogonal techniques to 'resolve' or separate the component proteins. During the first 'dimension,' IEF is used to separate the proteins by their charge. This is followed by denaturing SDS-PAGE for the second dimension, where proteins are separated on the basis of their size or MW. IEF is an electrophoretic method that separates proteins based on their respective pI values. Being amphoteric molecules, proteins carry a varying net charge, depending on the pH of their surroundings. This is the sum of all the negative and positive charges on its constituent amino acid side chains (including those of modified groups), as well as its amino (N) and carboxy (C) termini. During IEF, proteins move along a pH gradient under an applied electrical field (potential difference) until they reach a position where their net charge is zero. Thereby, proteins are effectively separated on the basis of their small intrinsic differences in charge. SDS-PAGE is an electrophoretic method whereby proteins are separated by their molecular weights during migration through the pores of a polyacrylamide gel matrix containing an ionic detergent, SDS. SDS molecules effectively mask the intrinsic charge of the proteins by coating their surfaces, forming anionic complexes that have an overall net charge which is roughly proportional to the molecular weight of the protein. This is generally performed under reducing conditions, e.g. a reducing agent such as dithiothreitol (DTT) is added, to cleave intra- and inter-protein disulfide bonds. This improves protein resolution and reduces the complexity of the pattern obtained.

After separation in the second dimension, protein components that are resolved on the polyacrylamide gel matrix can be detected and visualized by various chemical staining methods. The most common staining techniques involve the use of a highly-colored organic dye (Coomassie Brilliant Blue, known as Coomassie staining), silver ions (silver staining), or fluorescent dyes, such as Sypro® Ruby (Molecular Probes®) or Deep Purple™ (GE Healthcare). These three staining methods have differing protein detection sensitivities: Coomassie staining (detection limit ca. 8–16 ng), silver staining (detection limit ca. 2–4 ng) and fluorescent dyes (detection limits ca. 1–100 ng) (Poland et al., 2005). The choice of staining method depends upon the practical and experimental parameters (e.g. protein concentrations, cost, speed), as well as compatibility downstream applications, especially in the use of mass spectrometry (see below). A key parameter is dynamic range: the range of protein concentrations for which the intensity of the stain is directly proportional to levels of protein present in the spot. After the protein spots have been visualized by staining, gel scanning can be performed (using a flat-bed gel scanner/densitometer) to obtain a digitalized image for downstream analysis. In a typical set of experiments, several identical replicates are performed to obtain a 'reference map.' The image of this map may be systematically compared, i.e. 'matched' to gel images obtained using different experimental conditions, to identify changes in protein spot patterns and 'volumes' (spot intensity  $\times$  area). A variety of commercial 2-DE analysis software packages are available to perform such tasks in a highly reproducible and automated manner, e.g. ImageMaster (GE Healthcare), PDQuest (BioRad) and Progenesis (Nonlinear Dynamics) (Garrels, 1989; Appel et al., 1991; Rosengren et al., 2003; Medberry et al., 2005).

Despite its powerful ability to resolve complex protein mixtures, it should be noted that 2-DE was not initially widely used. Its greatest deficiency was its poor reproducibility. When this technology was first developed, the pH gradient that was used to separate proteins in the first dimension was often quite unstable and non-uniform in nature. At that time, protein IEF was performed using a tube-like gel containing a pH gradient that was constructed using soluble (freely-mobile) carrier-ampholytes. IEF gels containing a pH gradient that was composed of immobilized-ampholytes were subsequently developed (Bjellqvist et al., 1982). These were far more robust and uniform in nature, and are now widely used due to their excellent reproducibility. Both linear and non-linear pH gradients, which span a number of different pH ranges, are now commercially available (Bjellqvist et al., 1982; Cordwell et al., 2000; Görg et al., 2000; Patton, 2002).

In the early days of 2-DE, the other main limitation was the ability to identify the protein(s) present within the individual 'spot' on the gels. The Edman sequencing (degradation) approach, which was developed during the 1950s, was widely used for this purpose (Edman, 1950). This technique

involves chemically-labeling the amino acid residue at the N-terminus of the protein/peptide, which results in it being cleaved from the chain, such that it can be subsequently identified by a chromatographic or electrophoresis-based method. This process is repeated in an iterative manner, allowing the primary sequence of the protein chain to be identified in a stepwise fashion (Edman, 1950; Edman and Begg, 1967). However, Edman degradation is not applicable to proteins whose N terminal amino acid is post-translationally modified (e.g. N terminal formylation) and only a maximum of ca. 40 amino acids can be identified each time. When used for the identification of multiple protein spots on a 2-D gel, this method becomes increasingly tedious and expensive. In order for 2-DE to become a truly widespread approach for proteome analysis, an alternative method was required for the rapid, large scale identification of protein spots. In the 1990s this arrived, with the introduction of mass spectrometry (MS)-based approaches for protein identification.

### IDENTIFICATION OF PROTEIN SPOTS ON 2-D GELS USING MASS SPECTROMETRY

MS is a fast, sensitive and highly reproducible analytical technique that is used for the identification of chemical molecules, based on the determination of their molecular mass (Guerrera and Kleiner, 2005). It is widely used for the identification of unknown compounds, as well as the detection of compounds of known identity. Structural and compositional information may also be reliably inferred via analyzing the masses of breakdown products (molecular fragments) formed during mass spectrometric analysis. However, the analyte must be volatile and must be capable of being ionized, in order to be analyzed on a mass spectrometer (Aebersold and Mann, 2003). Progressive technical advances in MS approaches such as fast atom bombardment (FAB), matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI) have enabled increasingly larger peptides and then protein molecules to be analyzed (Morris et al., 1981; Karas and Hillenkamp, 1988; Fenn et al., 1989).

A mass spectrometer comprises three basic components: an ion source, a mass analyzer and an ion detector. The ionized peptide molecules are produced in the ion source, and then separated by their mass-to-charge ( $m/z$ ) ratios in the (electromagnetic) mass analyzer, before being detected and recorded. The output is in the form of a series of mass-to-charge ( $m/z$ ) peaks, known as a mass spectrum. Proteins (peptides) are identified by systematically comparing the  $m/z$  values obtained, which correspond to (singly or multiply-charged) ions of intact or fragmented molecules produced during the ionization process, with on-line or in-house databases using a variety of bioinformatics approaches (see below). Not only can MS be used to obtain the molecular mass of a protein, it can also provide compositional and

structural information. For example, post-translational modifications to the peptide backbone, such as glycosylation or phosphorylation can also be revealed (Aebersold and Goodlett, 2001). MS may be used in a high throughput manner for the accurate identification of protein components present within complex mixtures in combination with other techniques, making it an invaluable component of modern proteomic analysis.

MS is routinely combined with 2-DE for proteome characterization, as MS alone cannot be used to analyze, identify or quantify the protein components present in complex mixtures. Proteins isolated from spots on the gel are generally broken down into peptide fragments prior to MS analysis, as these molecules can be identified far more accurately, and more conveniently by this technique (Rabilloud et al., 2009). This is generally achieved by chemical fragmentation, e.g. treatment with cyanogen bromide (CNBr), or by using a non-specific or (partially) site-specific protease, also known as an endoprotease (e.g. trypsin, chymotrypsin, Lys-C, and Glu-C). Various forms of chromatography (e.g. reverse phase-high performance liquid chromatography, RP-HPLC) or electrophoresis (e.g. capillary electrophoresis, CE) are commonly employed to pre-fractionate and resolve peptide components prior to MS analysis. Both RP-HPLC and CE can be directly coupled 'in-line' to the mass spectrometer, allowing automation and increased convenience and throughput. Further integration with 2-D gel scanners and robotic spot pickers/excisers enables hundreds of protein species resolved on 2-D gels to be analyzed in a highly-automated and accurate manner (Houthaeve et al., 1997; Traini et al., 1998; Breen et al., 2000).

There are two main approaches for the identification of proteins using MS-based techniques: peptide mass fingerprinting (PMF) and peptide sequencing. PMF is a popular method for protein identification due to its efficiency and simplicity. The idea was first suggested by W. J. Henzel in 1989, and was introduced in 1993 (Yates et al., 1993; James et al., 1994). A peptide mass fingerprint corresponds to the set of masses of its component peptide fragments, which are generally produced by digestion with a sequence-specific protease. Trypsin is most commonly used, which cleaves the protein backbone at the carboxyl side of arginine or lysine residues. This generates a set of ca. 10–50 'tryptic peptides' for an average bacterial protein. Like a fingerprint, it is essentially unique for any given protein. In PMF, the set of masses experimentally determined for the peptides produced after protease digestion (e.g. tryptic peptides) is systematically compared to sets of peptide masses obtained after analogous *in silico* (virtual) digestion of thousands (or millions) of protein sequences held in the large protein databases. Different algorithms had been developed for the comparison of measured masses with theoretical masses for protein identification by means of a scoring system with compromising sensitivity and specificity. The most commonly used algorithms include

MASCOT, MS-Fit and SEQUEST (Eng et al., 1994; Clauser et al., 1999; Perkins et al., 1999). Although the algorithms are different, proteins are identified by analogous comparative methods, with a score assigned for each protein, based on how closely the experimental and theoretical peptide datasets match. A positive identification 'hit' is defined as a score above an arbitrary confidence threshold. This is generally achieved when a large number of individual peptide masses can be matched to the virtual protein digest with high levels of mass accuracy. The protein is deemed 'unidentified' when no high confidence hit is obtained, i.e. the score is below the threshold value (McHugh and Arthur, 2008). Various statistical or database searching 'filters' may be applied to maximize the probability of identifying the correct protein match, e.g. restricting the database search to proteins encoded by a particular species.

As peptide mass fingerprinting cannot be reliably used to analyze complex protein mixtures, it is most commonly used for the identification of relatively pure proteins, such as those present within individual gel spots obtained by 2-DE. When interfaced with 2-DE, a typical PMF procedure involves several steps: (1) excision of the spot of interest; (2) treatment with an appropriate protease (e.g. trypsin); (3) extraction (purification) of the peptide fragments produced; followed by (4) mass spectrometric analysis. These steps may be performed manually or with the aid of robotic systems. MALDI-MS is generally used for PMF, as it is relatively tolerant of contaminating salts and other impurities. A prerequisite for the success of protein identification via PMF is the existence of an identical or highly similar protein sequence (or its encoding DNA sequence) in a searchable database (e.g. the NCBI GenBank database). When studying multiple proteins from a single organism, which is typically the case for 2-DE experiments, PMF can only be reliably used when the full genome sequence is available. A notable disadvantage of PMF is that it does not establish the actual amino acid sequence of the peptide (protein fragment), which may result in ambiguities when peptides of identical (or near-equal) masses are identified. To circumvent this problem, a peptide sequencing-based approach has been devised for protein identification (Aebersold and Goodlett, 2001).

Sequence-based approaches for protein identification involve the analysis of (trypsin-generated) peptide fragments using tandem mass spectrometry (MS/MS). MS/MS involves two consecutive stages of mass determination. In the first step, individual peptide ions produced during the first MS procedure (i.e. after ionization, ion separation and ion detection) are fragmented into various component ions by processes such as collision-induced dissociation (CID). The electronic and vibrational excitement of the peptide bonds causes the peptide chain to be fragmented in a predictable manner. This generates predictable sets of fragment ions, which correspond to various forms of the component amino acid residues (Aebersold and Goodlett, 2001). These fragment ions

are then introduced into the second mass spectrometer, which accurately determines their individual  $m/z$  values. The generated 'ladder' of fragment ions apparent in the mass spectrum can be reliably 'read' to reveal the amino acid sequence of the peptide fragment. Mass spectral data is compared with the protein databases using various customized software packages to identify the protein. This approach is somewhat analogous to Edman degradation, with the notable difference in that tryptic peptide fragments correspond to positions found throughout the protein chain, not just at the N-terminus. MS/MS data can also be used for *de novo* peptide sequencing, as the 20 common amino acids all have unique masses (with the exception of leucine and isoleucine, which are isobaric, and are therefore often problematic). This approach is highly useful for the identification of proteins of unknown origin, or proteins from organisms whose genomes are yet to be sequenced. Amongst many other applications, it is also invaluable for the identification and characterization of proteins with extensive post-translational modifications, or proteins containing multiple amino acid polymorphisms compared to the genome-sequenced (reference) strains.

## ADVANCEMENTS AND CURRENT STATUS OF 2-DE

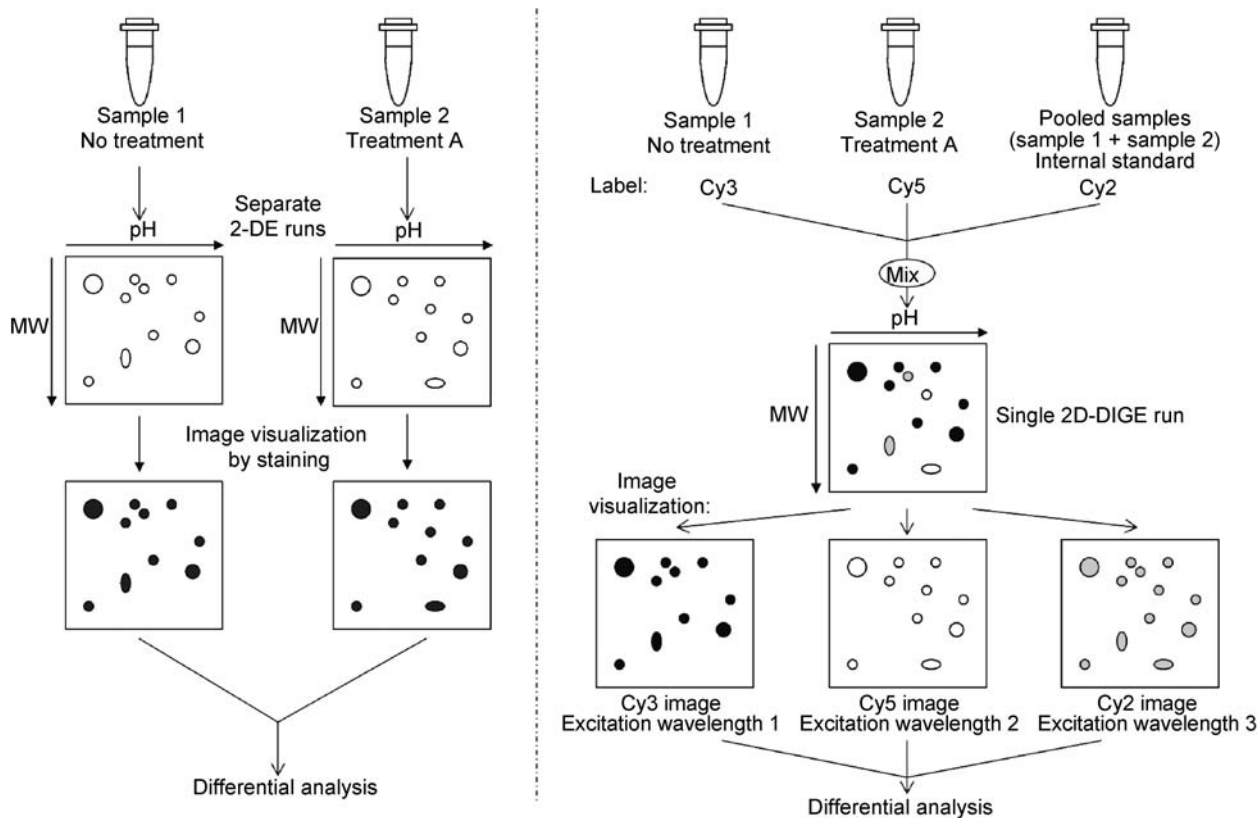
One of the most notable advancements in proteome analysis made over recent years has been the development of an approach known as 2-D differential gel electrophoresis (2D-DIGE) (Unlü et al., 1997). This technique is an invaluable approach for analyzing global changes in protein expression patterns within a certain system (e.g. a bacterial organism) that has been subjected to several different conditions. For example, it may be used to study protein expression patterns in bacterial cells that have been cultured under different conditions, or have been subjected to challenge with different chemicals or drugs. It involves the chemical labeling of protein samples with amine-reactive fluorescent dye molecules, e.g. N-hydroxysuccinimidyl derivatives of cyanine dyes (CyDyes: Cy2, Cy3, Cy5), prior to analysis using 2-DE. By using an orthogonal labeling approach, i.e. by covalently labeling two or three different protein samples with different fluorescent dyes that have distinct excitation and emission characteristics, they may be mixed and analyzed together on the same 2-D gel (Fig. 1) (Unlü et al., 1997; Marouga et al., 2005). Gels are visualized using a commercial fluorescent gel scanner/densitometer to quantify the fluorescence intensity, which is proportional to the amount of (labeled) protein present in each spot. An internal standard or control sample is usually included, to enable sample multiplexing. Gel to gel variation is reduced, as all samples are co-separated under identical electrophoretic conditions. Two matched sets of amine- or thiol-reactive CyDye fluorophores are available for minimal and saturation labeling, respectively. Minimal labeling involves adjusting the relative ratios of fluorescent labeling reagent and protein, so that an average of one lysine residue

on each protein molecule is covalently modified with one fluorescent dye molecule. This is the standard approach, which is used when protein samples are readily available (> ca. 50 ng). Minimal labeling can reduce the occurrence of protein precipitation, due to the relative hydrophobicity of the (Cyanine) fluorescent dye molecules. When the protein sample is scarce, saturation labeling would be preferred, as it can be reliably used to detect differences in protein abundance when only ca. 5 ng of protein sample is available. Saturation labeling involves the modification of all available cysteine residues, typically of fairly low abundance in most proteins, using a thiol-reactive fluorescent dye.

Compared with traditional colorimetric protein stains, fluorescent dyes have high detection sensitivities and a good dynamic range. 2D-DIGE is capable of quantifying small differences in protein levels, making it a popular approach for large-scale expression profiling in a range of biological systems. In the field of bacterial proteomics, 2D-DIGE has become a standard comparative proteomic approach for the study of gene regulation, stress responses, virulence mechanisms as well as environmental adaptation in both Gram-negative and Gram-positive species (Gade et al., 2005; Al Dahouk et al., 2008; Alteri et al., 2009; Franco et al., 2009). Investigators have previously compared directly the use of 2D-DIGE and colorimetric protein stains to quantify changes in protein expression patterns in *E. coli* cells treated with benzoic acid (Yan et al., 2002; Han and Lee, 2006). Analogous results were obtained, revealing that both techniques may be reliably used for protein expression profiling. However, it must be mentioned that there are situations where the use of 2D-DIGE may be less effective than conventional 2-DE approaches. Proteins with a low percentage of lysine residues may be 'under-represented' as they are labeled less efficiently than those with high lysine content. They may appear as protein spots of low intensity (abundance) in 2D-DIGE system because of its low lysine content (Marouga et al., 2005).

## LIMITATIONS OF 2-DE

Despite its impressive capabilities for protein separation and identification, 2-DE may not be a good approach for the analysis of hydrophobic proteins, proteins of low abundance, or samples that contain proteins with concentrations in a large dynamic range (Wilkins et al., 1998; Corthals et al., 2000; Rabilloud et al., 2010). For proteins to be resolved and appear as a single spot on a 2-D polyacrylamide gel, they should be completely soluble, denatured and reduced under the conditions employed. However, this is often highly problematic for certain types of proteins, due to their intrinsic chemical or biophysical properties. Hydrophobic proteins such as membrane-integral proteins or cell surface-associated proteins generally have poor solubility under the chemical conditions used during the IEF dimension, which is



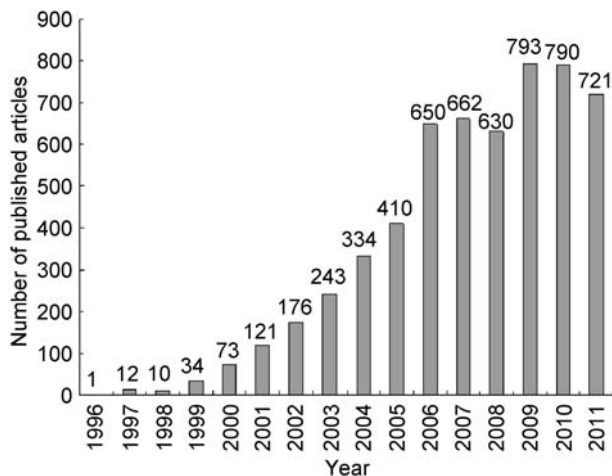
**Figure 1. Overview of variations in 2-DE technology in comparative proteomic analysis.** Left panel illustrates the flow of traditional 2-DE. Different protein samples with different treatments are subjected to separate 2-DE run to obtain the proteome profile of a particular sample. The images are visualized by staining with either silver, Coomassie blue or fluorescent stains. Comparative analysis is then performed by comparing the different 2-DE images. Right panel illustrates the flow of 2D-DIGE. Different protein samples with different treatments are pre-labeled with fluorescent dyes with different spectral characteristics prior to 2-DE. Samples of different sources are reacted with either Cy3 or Cy5 dyes. Pooled protein samples under investigation are labeled with Cy2 dye to serve as internal standards. All labeled samples are mixed and undergo 2-DE simultaneously on the same gel. Different protein samples resolved on the gel can be visualized using a fluorescent scanner that can excite different wavelengths, corresponding to the fluorescent dyes that are used to label the protein samples. Differential analysis can then be carried out to identify proteins of different expression levels. 2D-DIGE is superior over traditional 2-DE by reducing the number of gels for analysis and reducing experimental errors such as gel-to-gel variations.

performed in buffer of low ionic strength that does not contain ionic detergents. Various attempts have been used to improve their solubility, including the use of alternative detergents and chaotropes (denaturing agents). Even though some notable improvements have been achieved (see section below), hydrophobic proteins remain an intrinsic problem for 2-DE due to the IEF step. Low abundance proteins may be reliably visualized if, more than a few nanograms are present. However, they may be partially or completely occluded (masked) by high abundance spots with similar MW and pI values. Furthermore, it is very difficult to visualize and quantify proteins on the same gel when their relative concentrations differ by many orders of magnitude, as is the case for many biological samples, most notably blood plasma (Greenough et al., 2004). As a result, 2-DE may greatly under-represent the whole proteome. In such situations, strategies including the depletion of high abundance proteins, the use of very narrow-range IPG strips (~1 pH unit) or proteome

subfractionation (see section below), can be employed to enable the detection of low abundance proteins (Greenough et al., 2004; Chevalier, 2010; Rabilloud et al., 2010).

### WHY USE 2-DE?

2-DE is an extremely robust and reliable method for the simultaneous analysis of many hundreds of different protein species present within a single sample. Even though it was initially developed more than 35 years ago, it remains a highly used proteomic technique. To date, there are more than 5000 published articles reporting the use of gel-based approaches for analyzing protein patterns in a plethora of biological systems (Fig. 2). The reproducibility of the 2-DE approach has been investigated in several inter-laboratory studies, which revealed that when experiments were performed accurately, minimal variations were observed (Corbett et al., 1994; Blomberg et al., 1995). The most common source of variability



**Figure 2. Chronological list of articles published in proteomic field by means of two-dimensional electrophoresis technology as of Dec 2011.** The list is generated using PubMed search with keywords 'two-dimensional electrophoresis,' '2-DE,' '2-D electrophoresis,' '2-D gel electrophoresis,' '2D-PAGE,' '2D-DIGE' or 'DIGE' and each of 'proteome,' 'proteomic' or 'proteomics.' A growing trend is observed in the late 1990s to early 2000s. The number of publications is leveled in recent years, suggesting an importance of gel-based approaches in proteomics is still evident at present.

ty lies within the sample preparation step, as there can be multiple confounding factors. It is often difficult to rapidly stabilize or quench the highly-dynamic biological samples in a consistent manner, so that the entire protein content can be reliably extracted and purified from contaminating factors (e.g. salts that may interfere with the IEF step).

The detection of post-translationally modified proteins is another notable advantage of using a 2-DE approach. Based on the change in  $pI$  value, the modified and unmodified protein forms can be separated during the IEF dimension. This enables their identification, quantification and characterization during the second dimension and subsequent mass spectrometric analysis. Modifications including glycosylation, phosphorylation, or the removal of trypsin cleavage site by lysine acylation or methylation, result in changes in overall protein charge, or changes in peptide length or hydrophobicity. However, such approaches may be technically demanding (Sarioglu et al., 2000; Rabilloud et al., 2002; Wang et al., 2002).

## APPLICATIONS OF 2-DE IN BACTERIAL PROTEOMICS

Since its introduction, 2-DE has commonly been used to study bacterial proteomes, as these are generally considerably less complex than eukaryotic systems. Generating a reproducible pattern of separated proteins with high resolution, 2-DE offers the opportunity for a large-scale qualitative and

quantitative comparison of bacterial proteomes. It is a powerful tool for the delineation of bacterial proteomes expressed under different conditions, enabling these to be correlated to biological pathways or other aspects of cellular physiology (Henzel et al., 2003; Völker and Hecker, 2005). In combination with MS, and the plethora of bacterial whole-genome sequence data currently available, it is often possible to identify almost every single protein spot on the 2-D gel.

Microbial physiology is the field concerned with the study of the metabolic activities, subcellular structure, growth and survival of microorganisms. It has been viewed as a central discipline in microbiology (El-Sharoud and Rowbury, 2006). Over years, the use of 'systems biology' approaches to study or model the physiology of bacteria has gained increasing attention, as cellular functions are controlled in an integrated and coordinated manner. The use of proteomics, especially 2-DE, is of central importance to such efforts, as it can be used to take 'snapshots' of a large proportion of the entire protein complements of the cell. As transcriptomic (mRNA) information is not always capable of giving a true reflection of the adaptation of microbes towards their changing environment, proteomics is an invaluable tool for directly identifying the functions or connectivities of specific gene products (Han and Lee, 2006; Hecker et al., 2008; Bumann, 2010).

## Protein profiling of bacteria

Profiling the proteome of a bacterium grown under certain conditions can help us to establish a fundamental understanding of its cellular physiology. Since protein expression is an energetically costly process, generally speaking, only the specific proteins directly required for growth, survival or pathogenicity are expressed under a particular physiological condition. Consequently, cataloging all the proteins expressed under a specific set of conditions (including the identification of post-translational modifications) can help identify the components responsible for maintaining certain cellular activities. A bacterium, such as *E. coli* which has ca. 4300 genes, typically expresses 1000–1500 proteins under standard laboratory growth conditions. This is ideally suited to 2-DE, in which it is possible to simultaneously visualize and quantify a comparable number of protein spots on a single gel (Kalia and Gupta, 2005). Using 2-DE expression profiling under various culture conditions, it is often possible to identify sets of proteins involved in basic cellular metabolism *versus* those expressed in response to various cellular stresses or external stimuli. While the majority of 'house-keeping' proteins that function in central metabolic pathways are expressed at relatively constant levels during the active growth stage, many proteins involved in stress response or adaptation are significantly up-regulated from a lower 'basal' level under certain growth conditions. To further investigate the putative physiological roles or biological functions of proteins,

various fractionation approaches may be used to establish their predominant cellular localizations. In such approaches, whole cell proteomes may be compared with various 'sub-proteomes,' which may include extracellular proteins, membrane or cell wall associated proteins, DNA/RNA associated proteins, or cytoplasmic fractions.

### Whole cell proteome

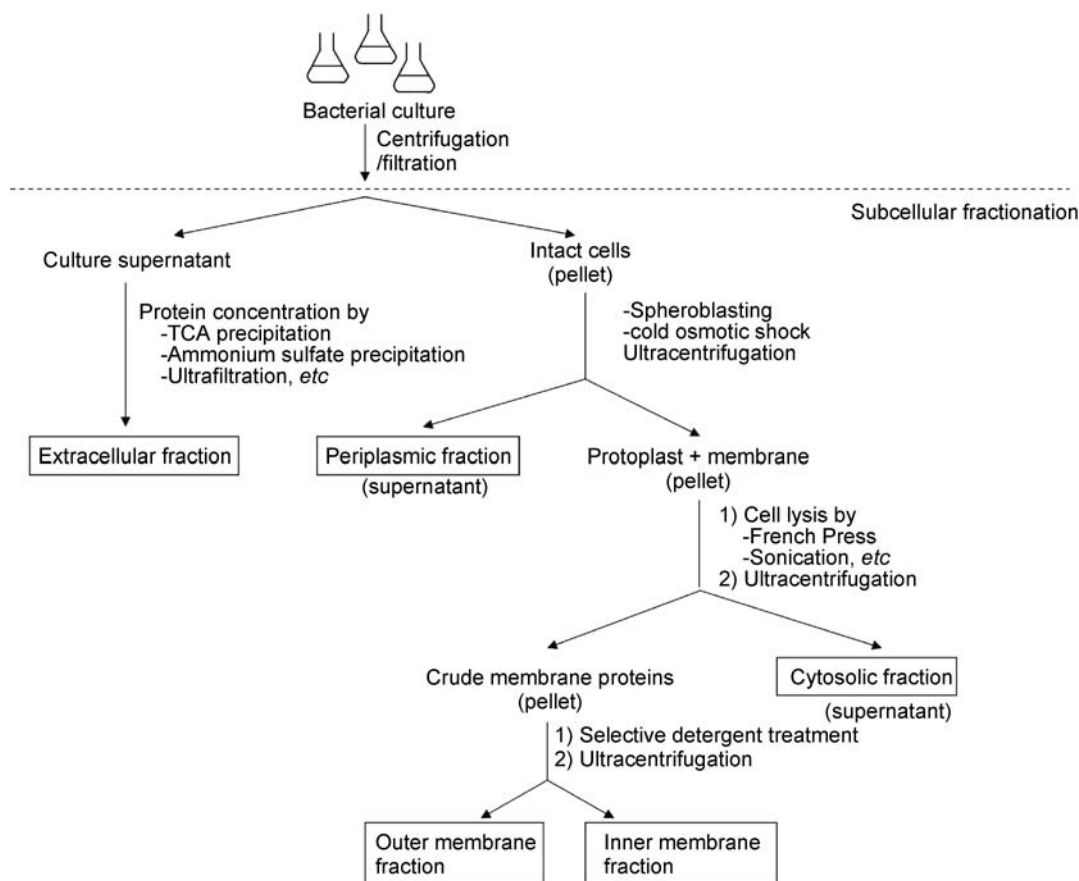
2-D gel image database (SWISS-2DPAGE) was established in 1993, with the premise of having an internet-accessible resource where all the spots on high quality 2-D gels of an organism's proteome were clearly annotated and identified (Appel et al., 1993). Before 2011, it was jointly maintained by the Biomedical Proteomics Research Group (BPRG) of Geneva University and the Proteome Informatics Group of the Swiss Institute of Bioinformatics (SIB), and is now maintained by Swiss-Prot and Vital-IT of the Swiss Institute of Bioinformatics. This database collates experimental data for each spot on the gel, including its identity, *pI* and *Mr* values, amino acid composition, peptide masses, and/or peptide sequences, as well as the corresponding bibliographical references, by integrating with other relevant databases such as SWISS-PROT. In this database, the majority of proteins were identified by microsequencing, immunoblotting, amino acid composition analysis, gel comparisons and PMF. 2-D gel 'reference maps' are currently available for 7 organisms, including *E. coli* and *Staphylococcus aureus* strain N315. Among the 36 reference maps available as of August 2011, 8 are dedicated to *E. coli*, which include a series of narrow range pH gradients as well as 2D-DIGE gels. More than 70% of the conserved *E. coli* proteome is represented in this resource. This is just a single resource; however, reference 2-D gel map has also been described for many other bacteria, especially pathogenic bacteria including: *Agrobacterium tumefaciens*, *Bacillus anthracis*, *Bacillus subtilis*, *Helicobacter pylori*, *Lactococcus lactis*, *Listeria monocytogenes*, *Mycoplasma pneumoniae*, *Salmonella enterica*, *Shigella flexneri*, methicillin-resistant and methicillin-sensitive strains of *Staphylococcus aureus*, *Streptococcus pneumoniae*, *Streptococcus thermophilus* and *Vibrio cholerae* (O'Connor et al., 1997; Anglade et al., 2000; Jungblut et al., 2000; Perrin et al., 2000; Regula et al., 2000; Cordwell et al., 2002; Liao et al., 2003; Coelho et al., 2004; Eymann et al., 2004; Folio et al., 2004; Rosen et al., 2004; Encheva et al., 2005; Wang et al., 2005; Encheva et al., 2006). Through concerted and exhaustive efforts, it is often possible to resolve and identify more than a thousand protein spots in a comprehensive 2-D reference map. As a result, many hundreds of proteins of both known (predicted) and unknown or uncharacterized function are annotated on these 2-D reference gels. For example, the 2-D proteome map for *Bacillus subtilis* grown in a defined minimal medium at mid-log phase was constructed for pro-

teins with *pI* values in the range 4–7 (Hecker and Völker, 2004). This enabled the identification of proteins involved in central metabolic pathways, including carbohydrate metabolism, the biosynthetic pathways of many amino acids, fatty acid and purine/pyrimidine metabolism, as well as proteins involved in RNA transcription, translation and DNA replication. As such, this 2-D gel provided a snapshot of the major proteins involved in basic cellular physiology of this bacterium. As was predicted, the majority of the most abundant proteins were those involved in central physiological processes, such as energy metabolism, protein synthesis and basic cellular processes (housekeeping proteins). This enabled the metabolic pathways that were actively working under the specific growth conditions to be reconstructed, based on the presence of the corresponding enzymes (Hecker and Völker, 2004).

2-DE may also be used to investigate the proteomes of uncultivable bacteria. The bacterial endosymbiont of the deep-sea tube worm *Riftia pachyptila*, a single species of sulfide-oxidizing  $\gamma$ -Proteobacteria, cannot be cultivated outside its host. One study has successfully established a reference 2-D gel of the soluble intracellular and membrane associated proteins of this bacterium without cultivation, by direct isolation from its host (Markert et al., 2007). It was demonstrated that the bacterium had a complex metabolism, with a substantial portion of proteins devoted to sulfide oxidation. The involvement of a reductive tricarboxylic acid cycle in energy metabolism was also established, which was not previously known to occur. This highlights the important role that experimental proteomics can play in understanding 'real' life processes.

### Subcellular proteome

For most bacteria, the majority of proteins present in the whole cell proteome, which generally refers to all the proteins that can be extracted after cell lysis, are cytosolic proteins. In the cytosolic fraction, housekeeping and structural proteins generally predominate. As the abundance of cytosolic proteins often varies by more than 4 or 5 orders of magnitude, low abundance proteins are commonly 'masked' by those of high abundance. Thus, the 2-DE analysis of the whole cell proteome does not often reveal a complete picture of the expressed protein complement of a bacterium, under the specific growth conditions employed (Ishihama et al., 2008). Sample pre-fractionation is therefore commonly utilized as a means to preferentially enrich a target protein group for analysis. This approach, also known as subcellular proteomics, examines the bacterial proteins present in a specific subcellular location (fraction), thus eliminating highly abundant proteins and reducing the overall complexity (Les-cuyer et al., 2004). In a bacterial cell, proteins are often targeted to certain locations (e.g. cell pole, nucleoid, midcell) in



**Figure 3. Subcellular fractionation of Gram-negative bacterial cell culture.** Centrifugation was utilized to separate intact bacterial cells from culture medium and extracellular proteins. Protein samples in the supernatant constitute the extracellular fraction. Periplasm can be isolated by cell spheroblasting or cold osmotic shock treatment followed by ultracentrifugation. The cytosolic fraction and membrane fraction can be isolated separately via various cell lysis procedures, detergent treatment and ultracentrifugation. The separation procedures are illustrated in the figure.

order to perform specific functions. Therefore, identifying the subcellular localization of an uncharacterized protein can help identify a putative function. In one 2-DE study, researchers fractionated the *E. coli* proteome, to characterize the protein composition of various cellular extracts (Lopez-Campistrous et al., 2005). They analyzed 2160 protein spots, identifying 575 unique protein species, including 151 hypothetical proteins, 76 proteins of unknown function, and 222 proteins not previously assigned in the SWISS-PROT database. This highlights the importance of sample fractionation in 2-DE. In general, proteins can be found in cytoplasm, inner membrane, outer membrane, periplasm (in the case of Gram-negative bacteria) or extracellularly (secreted). Various procedures can be used to obtain specific cell fractions, enabling their protein content to be determined (Fig. 3). For Gram-negative bacteria, the periplasmic extract can be isolated by cell spheroblasting, followed by differential centrifugation, or by a cold osmotic shock method. After separation of periplasmic proteins, cell lysis can be used to isolate the cytoplasmic proteins, with the membrane fraction separated by centrifu-

gation. The inner and outer membranes of Gram-negative bacteria can be isolated by selective cell lysis using lysozyme/EDTA treatment, by mechanical methods such as the use of a French Press, or by using commercially available chemical lysis reagents (mostly detergents) (Thein et al., 2010). After lysis, distinct inner and outer membrane vesicles are formed, which can be separated from each other by density centrifugation (with the inner membrane vesicles having a higher density). Selective detergent treatment, such as with Triton X-100, which preferentially dissolves cytoplasmic membranes, can also be used to separate inner and outer membrane protein fractions. In Gram-positive bacteria, cell surface proteins have been successfully isolated for analysis. The methodology that is known as surface shaving liberates surface exposed protein domains by digesting the cell surface with enzymes such as trypsin or proteinase K. Choosing the appropriate fractionation method is essential to the success of subcellular proteomics, as cross-contamination of the various fractions will seriously affect the results obtained.

### Secretome/exoproteome

In order to better understand how bacteria interact with, and be affected by the changes in their immediate environment, cell surface (cell wall or membrane associated) and extracellular proteins have increasingly being subjected to detailed proteomic investigation using 2-DE techniques. For pathogenic or commensal species of bacteria, these proteins may mediate key interactions with their proximal host cells, or components of the immune system. Consequently, these proteins may play key roles associated with disease pathogenesis (Jungblut, 2001; Tjalsma et al., 2004; Thein et al., 2010). Proteins secreted from pathogenic species of bacteria have a wide range of functions, acting as degradative enzymes (e.g. proteases) or virulence factors (e.g. toxins), and may be essential for cell viability and proliferation (e.g. providing nutrients). In Gram-negative and -positive species, secreted proteins are transported via a variety of secretion systems to various locations, including being anchored to the cytoplasmic membrane, associating with the cell wall, released into the extracellular milieu, or injected into a host cell (Desvaux et al., 2009). Proteins that are found in the extracellular milieu constitute the extracellular proteome, which is also known as the exoproteome or secretome. These proteins have been shown to contribute to bacterial adaptability to environmental changes and play crucial roles in infection (Tjalsma et al., 2004; Desvaux et al., 2010). Due to their key functions, extracellular proteins are generally good targets for antimicrobial therapies. Furthermore, the immunogenic properties of many proteins secreted by pathogenic bacteria make them prospective candidates for the development of subunit- or epitope-based vaccines (Bumann et al., 2004; Gupta et al., 2009).

With ongoing advancements in computational technologies and genome sequencing efforts, secreted bacterial proteins can be predicted or identified based on the secretion signals present in the encoded protein sequence (Desvaux et al., 2009). However, *in silico* predictive approaches are generally based on previously reported experimental results, and are often error prone (Antelmann et al., 2001; Gardy and Brinkman, 2006). In addition, several secretion signal independent secretion pathways have been identified (Bendtsen et al., 2005). Proteomics therefore represents the most straightforward way to directly identify extracellular proteins *in vivo*. To study the bacterial exoproteome, the protein content of the cell culture supernatant is generally collected either by protein precipitation, ultrafiltration or a salting-out method, then analyzed by 2-DE. It has successfully contributed to the delineation of several secretory pathways in bacteria. One notable example is the elucidation of the Twin arginine translocation (Tat) pathway in *B. subtilis*. The discovery and experimental validation of the first Tat-substrate, secreted alkaline phosphodiesterase (PhoD), in *B. subtilis* was achieved through the use of 2-DE analysis

(Jongbloed et al., 2000). The secretomes of multiple *tat* mutant strains were compared with wild type *B. subtilis*, which revealed that proteins that were secreted by the Tat pathway were not solely dependent on the presence of RR- or KR-motif, the typical twin-arginine signal peptide in Gram-negative bacteria (Jongbloed et al., 2000).

Secretome analysis of various pathogenic bacteria, including *Aeromonas hydrophila*, *L. monocytogenes*, *H. pylori*, *B. anthracis*, *B. subtilis*, *S. aureus* and *Pseudomonas aeruginosa*, has not only identified known virulence factors, but has also revealed previously uncharacterized secreted proteins that were involved in pathogenesis (Bumann et al., 2002; Nouwens et al., 2002; Tjalsma et al., 2004; Trost et al., 2005; Chittlaru et al., 2006; Sibbald et al., 2006; Yu et al., 2007). For example, in *S. aureus*, the function of the diacylglycerol transferase (Lgt) virulence factor was revealed by comparing the exoproteomes of wild type and *Lgt* mutants with impaired virulence (Sibbald et al., 2006). In the fish and human pathogen *A. hydrophila*, 2-DE was used to identify a variety of virulence factors, including lateral flagellar proteins, proteases, lipases, hemolysin, aerolysin and glycerophospholipid cholesterol acyltransferase that were secreted to the cell surface, or released into the media (Yu et al., 2007). In *Legionella pneumophila* (the bacterial pathogen responsible for legionnaire's disease), type II protein secretion is known to be involved in virulence. DebRoy et al. used a comparative 2-DE analysis of the culture supernatant of wild type *L. pneumophila* versus a type-II secretion mutant, to identify 20 proteins via the type-II secretion pathway (DebRoy et al., 2006). This included a previously undescribed chitinase A protein, which was subsequently implicated in infection using an A/J mouse model, promoting persistence in the lung. These examples highlight how 2-DE may make important contributions towards understanding the molecular mechanisms of various extracellular proteins in bacterial virulence.

Cytosolic proteins are often found in the bacterial exoproteome. Some of these are classified as 'moonlighting' proteins. The term moonlighting means that, in addition to its well established activities, a protein also has an additional, but often completely unrelated, function (Henderson and Martin, 2011). A classic example is glyceraldehyde-3-phosphate dehydrogenase (GAPDH). Well-established as an important enzyme in glycolysis, with a typical cytoplasmic localization, GAPDH has also been found on the cell surface. First reported in group A streptococci, streptococcal GAPDH proteins are also located on the cell surface, where they bind to lysozyme, cytoskeletal proteins and fibronectin (Pancholi and Fischetti, 1992; Jeffery, 2009; Henderson and Martin, 2011). In addition to moonlighting proteins, contamination with cytosolic proteins released after cell lysis cannot be ignored. Therefore, a thorough analysis of the culture supernatant must be performed, such as monitoring for the presence of confirmed cytoplasmic protein markers.

### Membrane proteins

In addition to the secretome, membrane and cell wall-associated proteins are other classes of bacterial proteins that come into direct contact with the extracellular environment. As mentioned earlier, 2-DE has notable limitations for the analysis and resolution of membrane proteins. Nonetheless, extensive efforts have been devoted to devising methods that can be used to solubilize membrane proteins using combination of surfactants and chaotropes, enabling their analysis using 2-DE approaches (Santoni et al., 2000; Rabilloud, 2009). In their pioneering work, Molloy et al. described the use of an alkaline solution of sodium carbonate to dissolve outer membrane proteins (Molloy et al., 2000). Subsequent studies on the *E. coli* membrane proteins have optimized this sample preparation method, enabling 394 membrane proteins to be identified (Fountoulakis and Gasser, 2003). Following this milestone in examining bacterial membrane proteins using 2-DE, many Gram-positive and Gram-negative bacterial membranes proteomes, such as those from *Mycobacterium tuberculosis*, *Campylobacter jejuni*, *P. aeruginosa* and *S. aureus* have been investigated (Nouwens et al., 2002; Bunai and Yamane, 2005; Gatlin et al., 2006; Poetsch and Wolters, 2008). Membrane proteins are often highly immunogenic, and have been investigated in search of potential vaccine targets. They are usually analyzed in combination with immunogenicity studies, forming a field known as immunoproteomics (See section below).

### Differential proteomics

As mentioned earlier, the use of 2-DE to catalogue all the proteins expressed by a bacterium during a particular set of (culture) conditions may be extremely useful for gaining insight into its cellular physiology. However, comparative bacterial proteomics can be performed using several different approaches, including genotypic alternation, genetic differences, environmental perpetuation or differences in growth condition. In such approaches, various cellular processes are disrupted or perpetuated, in order to identify complex, interconnected regulatory or functional protein networks within the bacterium. In genotypic alteration, the global effects of a particular gene or group of genes (e.g. operon) are examined by comparing the protein profiles of wild type and mutant strains that are grown under identical conditions. Changes in the expression of all the proteins that are involved in the regulation reflect the regulatory circuits of the gene/operon under study, i.e. the 'regulon.' To analyze the effects of genetic differences, the proteomes of strains belonging to the same species with notably different phenotypes are systematically compared, to help identify the proteins underlying the phenotypic differences. For example, Franco et al. used 2D-DIGE to compare the expression of cytoplasmic and membrane proteins within carcinogenic and non-carcinogenic

strains of *H. pylori*, in an effort to reveal their carcinogenic potential. They identified 26 differentially expressed proteins, many of which had not previously been reported to be related to gastric cancer (Franco et al., 2009). A missense mutation of the flagellar protein FlaA, which led to the formation of a different FlaA isoform, was found to be related to the morphology of the carcinogenic phenotype.

2-DE is also an invaluable tool for understanding how bacteria adapt their physiology in response to various environmental conditions. The sophisticated networks underlying bacterial adaptation or survival processes under different conditions are reflected in cellular protein composition. Gaining a detailed knowledge of how pathogenic bacteria adapt to the changing environments encountered during their normal life cycle, including their interactions with natural and opportunistic hosts, is of fundamental importance. 2-DE proteomic techniques have enabled the downstream proteins controlled by key regulatory proteins to be identified, giving us significant insight into various cellular metabolic and regulatory networks. They have also enabled the identification of various stress-associated or stress-specific protein markers (Phillips and Bogoy, 2005). Environmental stresses influence a plethora of physiological activities in the bacterium. Protein expression or activities are systematically altered as a means to resist, or adapt to external stress. Over recent years, proteomic approaches have become increasingly popular for the characterization of changes in protein synthesis, after the exposure of the bacterium to various types of cellular stress. 2-DE has been the method of choice for research in the field of stress response/environmental adaptation, due to its capability of quantifying differential protein expression in a global manner (Renzone et al., 2005; Hecker et al., 2008). Extensive comparative proteomic analyses have been performed in *E. coli* and several other bacterial species, successfully revealing several novel regulatory networks (Han and Lee, 2006). The ability of a bacterium to detect and appropriately respond to various stresses/environmental stimuli is of paramount significance to its survival abilities within different hosts or natural reservoirs. Many bacterial pathogens sense certain environmental indicators such as the host's body temperature, osmolarity, pH, iron levels, as well as other host derived signaling molecules, and use these as cues for large-scale physiological adaptation (Hueck, 1998; Konkel and Tilly, 2000). Bacteria have also developed protective mechanisms, which respond to various chemical and physical stresses, such as heat, oxidative agents, UV radiation and heavy metals, in order to improve their survival ability (Hecker et al., 2008; Lau et al., 2011). To illustrate how 2-DE can contribute to the understanding of bacterial environmental adaptation, we highlight several regulatory pathways initiated in response to temperature changes.

Temperature is a critical environmental signal that varies dramatically as the bacterium inhabits different environments. Pathogens such as *S. flexneri*, *Bordetella pertussis* and

*Yersinia pestis* respond to temperature alterations by increasing the expression of sets of temperature-regulated genes during the entire period when the bacteria remain at an elevated temperature (Konkel and Tilly, 2000; Coote, 2001; Marceau, 2005). Temperature adaptation by pathogens differs from stress tolerance responses, such as a 'heat-shock' or 'cold-shock,' in which stress response is a transient process that triggers regulons which govern cellular damage control processes. It is a process in which bacteria respond and adapt to temperature changes in the extracellular environment by expressing certain subsets of genes that are required for energy conservation or the optimization of various physiological or growth processes (Konkel and Tilly, 2000). These mechanisms involve preserving the flexibility, topology and interactions of macromolecules such as DNA, RNA and proteins, maintaining the viscosity of the cell membrane, protection from cell disruption by producing osmoprotectants, and maintaining appropriate intracellular diffusion rates and enzyme kinetics (Thieringer et al., 1998; Phadtare et al., 1999; Gao et al., 2006). For example, protein expression in Gram-negative bacterium *Laribacter hongkongensis* cultured at 20°C and 37°C were compared using 2-DE (Woo et al., 2009). *L. hongkongensis* is a recently discovered  $\beta$ -Proteobacterium that is associated with community-acquired gastroenteritis and traveler's diarrhea (Yuen et al., 2001; Woo et al., 2004; Woo et al., 2005). It can inhabit in the intestines of freshwater fish, frogs and humans, and can also grow independently in freshwater environments (Woo et al., 2004; Teng et al., 2005; Lau et al., 2007; Ni et al., 2007; Lau et al., 2009). Therefore it needs to cope with varying environmental temperatures in these different growth habitats. Woo et al. identified 12 protein spots that were differentially expressed at these distinct environmental and human host temperatures. Among these, two isozymes of *N*-acetylglutamate kinase, a key enzyme in the arginine biosynthesis pathway, gave opposing expression patterns at these two temperatures (Woo et al., 2009). These 2-DE analyses clearly illustrated that temperature significantly altered the global expression of proteins, promoting the survival of the bacterium in these distinct niches. In *Yersinia enterocolitica*, the production and secretion of the Yop effector virulence factors are promoted at 37°C but not 30°C, whereas other virulence factors, including urease and the O-antigen outer core lipopolysaccharide are upregulated at lower temperatures (de Koning-Ward and Robins-Browne, 1997; Skurnik et al., 1999; Trülsch et al., 2003; Marceau, 2005).

Mimicking growth conditions in the host is also a way to investigate bacterial metabolism during infection processes, which is of key importance for understanding disease pathogenesis and proliferation. 2D-DIGE experiments revealed that uropathogenic *E. coli* (UPEC) grown in human urine had 84 differentially expressed proteins, compared with growth in LB media (Alteri et al., 2009). It was found that enzymes involved in peptide uptake, gluconeogenesis and

the tricarboxylic acid cycle were required for the 'fitness' of *E. coli* during urinary tract infection, suggesting that peptides and amino acids act as the primary carbon sources during infection of the urinary tract.

### Post-translational modifications (PTMs)

Although PTMs are traditionally assumed to predominate in eukaryotes, the importance of protein PTMs in bacterial systems has gained increasing attention over recent years (Deutscher and Saier, 2005). Proteins can be modified post-translationally in order to be activated, deactivated or to 'mark' them for degradation. In bacteria, protein modifications are involved in signal transduction and regulatory processes, modulating protein stability, activity, molecular interactions, as well as their subcellular localization. 2-DE is a useful approach for the identification of global PTM patterns, as they can be visualized directly on the 2-D gel. One of the most common PTMs is protein phosphorylation, in which the modified protein will have an overall net gain in negative charge, originating from the addition of a phosphate moiety to the reactive groups of various amino acid side chains. These commonly include hydroxyl groups (Ser, Thr, and Tyr), carboxyl groups (Asp), as well as amino or imidazole ring nitrogens (Lys, Arg, and His). In bacterial signaling, His and Asp phosphorylation is the most commonly studied regulatory mechanism in two component systems (Swanson et al., 1994). However, this type of modification is chemically unstable at neutral or lower pH values (pH < 8), and thus detection by MS is very difficult. There is now a growing attention towards the analysis of serine/threonine/tyrosine phosphorylation in bacteria, as it has been shown to be related to various key cellular processes including cell cycle, cell differentiation, stress response, and virulence and persistence (Deutscher and Saier, 2005). This mode of phosphorylation is acid stable, and therefore it is comparatively more practical for proteomic analysis. Using colorimetric Pro-Q Diamond® (Invitrogen) staining,  $^{32}\text{P}/^{33}\text{P}$ -radiolabeling or with the aid of phosphoprotein specific antibodies, phosphoproteins resolved on the 2-D gel can be revealed. 2-DE methods have been used to analyze the phosphoproteomes of several bacterial species, such as *E. coli*, *Corynebacterium glutamicum*, *Neisseria meningitidis* and *B. subtilis* (Cortay et al., 1986; Bendt et al., 2003; Macek et al., 2007; Bernardini et al., 2011). The phosphorylated proteins identified were involved in a variety of different cellular processes, such as central metabolism, with many having housekeeping functions. However, 2-DE may be less capable of detection of proteins involved in regulatory signal transduction pathways, as these proteins often exist in low abundance. In addition, the low recovery of (phosphorylated) peptides using in-gel digestion, usually leads to poor protein sequence coverage, which is the major problem associated with the detection of phosphorylation sites. Therefore, gel-free approaches have also been in-

creasingly adopted for the investigation of the bacterial phosphoproteome. Nevertheless, 2-DE techniques remain highly valuable and informative approaches for the characterization of bacterial protein phosphorylation patterns.

### Immunoproteomics

The use of 2-DE in conjunction with immunoblotting has proven to be a highly useful approach for the identification of immunogenic proteins. It has been particularly well used for investigations into pathogenic bacterial species, as a method to identify or characterize virulence factors, drug/vaccine candidates, for immunodiagnosis, as well as for protein PTM detection as described above (Cash, 2011). 2-D immunoblotting approaches are analogous to the procedures used for traditional 1-D gel immunoblotting (e.g. Western blotting). This involves electrotransferring the proteins from 2-D gel onto a chemically resilient (e.g. PVDF) membrane, sequential blotting with the primary then (conjugated) secondary antibodies, with subsequent detection of the label on secondary antibodies, to enable protein quantitation. The high resolving power of 2-DE enables the separation of complex protein mixtures, so that immunodetection should have good specificity, although co-migrating proteins can sometimes occur. Therefore, the detection result must be assessed with care. This type of 2-D immunoproteomic approach is used to identify immunogenic proteins (antigens) in pathogens as well as helpful for the understanding of the virulence factors and etiology. One example in the immunoproteomic analysis of *H. pylori* using sera from patients infected with *H. pylori* identified 32 protein candidates, in which 9 of them were novel antigens, including hypothetical protein HP0231, serine protease HtrA and Cag3 (Haas et al., 2002). Some of these novel antigens can trigger highly specific protective efficacies for *Helicobacter* in mouse models, showing that this approach can aid in diagnostic and vaccine design (Sabarth et al., 2002). Identification of immunogenic proteins can also be combined with subcellular protein fractions to provide an extensive picture of a pathogen's antigenic candidates. Immunoproteomic profiling of secreted and subcellular protein fractions of *Mycobacterium immunogenum* identified 33 protein species over the 4 cellular locations under investigation (Gupta et al., 2009). Apart from previously known mycobacterial antigens, 25 novel immunoreactive proteins are predicted to function in virulence as well as play diverse metabolic roles. The protein spots identified with major intensities were found to induce host response, demonstrating that they can be potential drug targets. Positive recognition of the protein species resolved on 2-D gel by patient or animal sera demonstrates the *in vivo* expression of the specific protein (antigen) and is deemed to be interacting with the immune system. In this regard, immunoproteomic approaches are of notable importance for vaccine development as well as the clinical diagnosis of bacterial infections.

### COMPLEMENTING 2-DE: THE EMERGENCE OF NON-GEL BASED APPROACHES

Over the past few years, a number of alternative techniques have been developed to overcome some of the limitations associated with gel-based methods for proteome analysis, especially regarding speed and sensitivity (Gevaert et al., 2007). Using a standard 2-DE approach, protein components are resolved as individual 'spots' on a gel, which are then picked, digested, and analyzed by MS. Recent efforts have focused on devising gel-free methods for protein separation, which can be efficiently coupled with existing MS methods for downstream analysis and identification (Aebersold and Mann, 2003; Gevaert et al., 2007; Dowell et al., 2008). In general, these gel-free methods involve 'shotgun proteomic' approaches, in which the complex protein mixtures are directly digested into many thousands of peptide fragments, which are then pre-fractionated using one or more chromatographic steps, followed by in-line MS analysis. Typically, proteins are trypsin digested, after which the resultant peptides are separated using anion or cation-exchange chromatography, and/or reverse phase chromatography (i.e. one or two dimension separations). Ion exchange (IEX) chromatography separates the peptides by charge, whilst reverse-phase (RP) chromatography separates them by hydrophobicity, resolving by means of two different biophysical properties (two dimensions). The use of liquid chromatography (LC) enables the convenient removal of common contaminants such as salt and detergents prior to MS analysis. The other major advantage with LC is that it may be directly coupled to a mass spectrometer, typically a tandem mass spectrometer, i.e. (LC-MS/MS), for greatly improved speed and efficiency. The high throughput and sensitive nature of this technique has made it an increasingly popular approach for (comparative) proteome profiling (Link et al., 1999; Qian et al., 2006).

Multidimensional protein identification technology (MudPIT) is another popular method for protein identification in very complex protein mixtures. It is based on the coupling of a two-dimensional LC separation to tandem mass spectrometry. In this method, the trypsin digested peptide mixture is loaded onto a biphasic microcapillary column containing a strong-cation exchange (SCX) resin upstream of an RP resin. Peptides are readily eluted from the RP resin under an acetonitrile gradient for subsequent online analysis by MS/MS (Washburn et al., 2001). A large number of peptides can be resolved, analyzed and identified in a single run, upon iterating the separation and elution steps using slightly different buffer/solvent parameters. Other non-gel proteomic approaches involving the multiplexing of different separation methods to mass spectrometry have also emerged. They have been used for high throughput protein identification, differential protein expression analysis, as well as PTM analysis. These methods have been comprehensively reviewed elsewhere (Roe and Griffin, 2006; Gevaert et al., 2007).

## CONCLUDING REMARKS AND FUTURE PERSPECTIVES

2-DE has contributed greatly towards our general understanding of bacterial proteomics. Its robust and reliable nature has made it a popular and enduring analytical technique. Despite its intrinsic limitations, it remains an extremely powerful and straightforward method for the global analysis of bacterial protein expression patterns. We envision that it will continue to evolve and mature, with fresh developments extending its resolution and range. Over forthcoming years both 2-DE and gel-free proteomic methods will be used in a highly complementary manner, to generate fresh insight into key roles played by proteins in bacterial biology, cellular physiology and disease pathogenesis.

## ACKNOWLEDGEMENTS

This work was partly supported by the Research Fund for the Control of Infectious Diseases (Commissioned study) of the Health, Welfare and Food Bureau of the Hong Kong SAR Government, Research Grant Council Grant, and University Development Fund.

## ABBREVIATIONS

2D-DIGE, 2-D differential gel electrophoresis; 2-DE, two-dimensional gel electrophoresis; MS, mass spectrometry; PMF, peptide mass fingerprinting; PTMs, post-translational modifications; SDS-PAGE, sodium-dodecyl sulfate polyacrylamide gel electrophoresis

## REFERENCES

- Aebersold, R., and Goodlett, D.R. (2001). Mass spectrometry in proteomics. *Chem Rev* 101, 269–295.
- Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. *Nature* 422, 198–207.
- Al Dahouk, S., Jubier-Maurin, V., Scholz, H.C., Tomaso, H., Karges, W., Neubauer, H., and Köhler, S. (2008). Quantitative analysis of the intramacrophagic *Brucella suis* proteome reveals metabolic adaptation to late stage of cellular infection. *Proteomics* 8, 3862–3870.
- Altarriba, M., Merino, S., Gavín, R., Canals, R., Rabaan, A., Shaw, J.G., and Tomás, J.M. (2003). A polar flagella operon (*flg*) of *Aeromonas hydrophila* contains genes required for lateral flagella expression. *Microb Pathog* 34, 249–259.
- Alteri, C.J., Smith, S.N., and Mobley, H.L. (2009). Fitness of *Escherichia coli* during urinary tract infection requires gluconeogenesis and the TCA cycle. *PLoS Pathog* 5, e1000448.
- Anglade, P., Demey, E., Labas, V., Le Caer, J.P., and Chich, J.F. (2000). Towards a proteomic map of *Lactococcus lactis* NCDO 763. *Electrophoresis* 21, 2546–2549.
- Antelmann, H., Tjalsma, H., Voigt, B., Ohlmeier, S., Bron, S., van Dijl, J.M., and Hecker, M. (2001). A proteomic view on genome-based signal peptide predictions. *Genome Res* 11, 1484–1502.
- Appel, R.D., Hochstrasser, D.F., Funk, M., Vargas, J.R., Pellegrini, C., Muller, A.F., and Scherrer, J.R. (1991). The MELANIE project: from a biopsy to automatic protein map interpretation by computer. *Electrophoresis* 12, 722–735.
- Appel, R.D., Sanchez, J.C., Bairoch, A., Golaz, O., Miu, M., Vargas, J.R., and Hochstrasser, D.F. (1993). SWISS-2DPAGE: a database of two-dimensional gel electrophoresis images. *Electrophoresis* 14, 1232–1238.
- Bendt, A.K., Burkovski, A., Schaffer, S., Bott, M., Farwick, M., Hermann, T., Farwick, M., and Hermann, T. (2003). Towards a phosphoproteome map of *Corynebacterium glutamicum*. *Proteomics* 3, 1637–1646.
- Bendtsen, J.D., Kiemer, L., Fausbøll, A., and Brunak, S. (2005). Non-classical protein secretion in bacteria. *BMC Microbiol* 5, 58.
- Bernardini, G., Laschi, M., Serchi, T., Arena, S., D'Ambrosio, C., Braconi, D., Scaloni, A., and Santucci, A. (2011). Mapping phosphoproteins in *Neisseria meningitidis* serogroup A. *Proteomics* 11, 1351–1358.
- Bjellqvist, B., Ek, K., Righetti, P.G., Gianazza, E., Görg, A., Westermeier, R., and Postel, W. (1982). Isoelectric focusing in immobilized pH gradients: principle, methodology and some applications. *J Biochem Biophys Methods* 6, 317–339.
- Blomberg, A., Blomberg, L., Norbeck, J., Fey, S.J., Larsen, P.M., Larsen, M., Roepstorff, P., Degand, H., Boutry, M., Posch, A., et al. (1995). Interlaboratory reproducibility of yeast protein patterns analyzed by immobilized pH gradient two-dimensional gel electrophoresis. *Electrophoresis* 16, 1935–1945.
- Breen, E.J., Hopwood, F.G., Williams, K.L., and Wilkins, M.R. (2000). Automatic poisson peak harvesting for high throughput protein identification. *Electrophoresis* 21, 2243–2251.
- Bumann, D. (2010). Pathogen proteomes during infection: A basis for infection research and novel control strategies. *J Proteomics* 73, 2267–2276.
- Bumann, D., Aksu, S., Wendland, M., Janek, K., Zimny-Arndt, U., Sabarth, N., Meyer, T.F., and Jungblut, P.R. (2002). Proteome analysis of secreted proteins of the gastric pathogen *Helicobacter pylori*. *Infect Immun* 70, 3396–3403.
- Bumann, D., Jungblut, P.R., and Meyer, T.F. (2004). *Helicobacter pylori* vaccine development based on combined subproteome analysis. *Proteomics* 4, 2843–2848.
- Bunai, K., and Yamane, K. (2005). Effectiveness and limitation of two-dimensional gel electrophoresis in bacterial membrane protein proteomics and perspectives. *J Chromatogr B Analyt Technol Biomed Life Sci* 815, 227–236.
- Cash, P. (2011). Investigating pathogen biology at the level of the proteome. *Proteomics* 11, 3190–3202.
- Chevalier, F. (2010). Highlights on the capacities of “Gel-based” proteomics. *Proteome Sci* 8, 23.
- Chittaru, T., Gat, O., Gozlan, Y., Ariel, N., and Shafferman, A. (2006). Differential proteomic analysis of the *Bacillus anthracis* secretome: distinct plasmid and chromosome CO<sub>2</sub>-dependent cross talk mechanisms modulate extracellular proteolytic activities. *J Bacteriol* 188, 3551–3571.
- Clauser, K.R., Baker, P., and Burlingame, A.L. (1999). Role of accurate mass measurement (+/- 10 ppm) in protein identification strategies employing MS or MS/MS and database searching. *Anal Chem* 71, 2871–2882.
- Coelho, A., de Oliveira Santos, E., Faria, M.L., de Carvalho, D.P., Soares, M.R., von Kruger, W.M., and Bisch, P.M. (2004). A pro-

- teome reference map for *Vibrio cholerae* El Tor. *Proteomics* 4, 1491–1504.
- Coote, J.G. (2001). Environmental sensing mechanisms in *Bordetella*. *Adv Microb Physiol* 44, 141–181.
- Coquet, L., Cosette, P., Dé, E., Galas, L., Vaudry, H., Rihouey, C., Lerouge, P., Junter, G.A., and Jouenne, T. (2005). Immobilization induces alterations in the outer membrane protein pattern of *Yersinia ruckeri*. *J Proteome Res* 4, 1988–1998.
- Corbett, J.M., Dunn, M.J., Posch, A., and Görg, A. (1994). Positional reproducibility of protein spots in two-dimensional polyacrylamide gel electrophoresis using immobilised pH gradient isoelectric focusing in the first dimension: an interlaboratory comparison. *Electrophoresis* 15, 1205–1211.
- Cordwell, S.J., Larsen, M.R., Cole, R.T., and Walsh, B.J. (2002). Comparative proteomics of *Staphylococcus aureus* and the response of methicillin-resistant and methicillin-sensitive strains to Triton X-100. *Microbiology* 148, 2765–2781.
- Cordwell, S.J., Nouwens, A.S., Verrills, N.M., Basseal, D.J., and Walsh, B.J. (2000). Subproteomics based upon protein cellular location and relative solubilities in conjunction with composite two-dimensional electrophoresis gels. *Electrophoresis* 21, 1094–1103.
- Cortay, J.C., Rieul, C., Duclos, B., and Cozzone, A.J. (1986). Characterization of the phosphoproteins of *Escherichia coli* cells by electrophoretic analysis. *Eur J Biochem* 159, 227–237.
- Corthals, G.L., Wasinger, V.C., Hochstrasser, D.F., and Sanchez, J.C. (2000). The dynamic range of protein expression: a challenge for proteomic research. *Electrophoresis* 21, 1104–1115.
- de Koning-Ward, T.F., and Robins-Browne, R.M. (1997). A novel mechanism of urease regulation in *Yersinia enterocolitica*. *FEMS Microbiol Lett* 147, 221–226.
- DebRoy, S., Dao, J., Söderberg, M., Rossier, O., and Cianciotto, N.P. (2006). *Legionella pneumophila* type II secretome reveals unique exoproteins and a chitinase that promotes bacterial persistence in the lung. *Proc Natl Acad Sci U S A* 103, 19146–19151.
- Desvaux, M., Dumas, E., Chafsey, I., Chambon, C., and Hébraud, M. (2010). Comprehensive appraisal of the extracellular proteins from a monoderm bacterium: theoretical and empirical exoproteomes of *Listeria monocytogenes* EGD-e by secretomics. *J Proteome Res* 9, 5076–5092.
- Desvaux, M., Hébraud, M., Talon, R., and Henderson, I.R. (2009). Secretion and subcellular localizations of bacterial proteins: a semantic awareness issue. *Trends Microbiol* 17, 139–145.
- Deutscher, J., and Saier, M.H. Jr. (2005). Ser/Thr/Tyr protein phosphorylation in bacteria - for long time neglected, now well established. *J Mol Microbiol Biotechnol* 9, 125–131.
- Dowell, J.A., Frost, D.C., Zhang, J., and Li, L. (2008). Comparison of two-dimensional fractionation techniques for shotgun proteomics. *Anal Chem* 80, 6715–6723.
- Edman, P. (1950). Method for determination of the amino acid sequence in peptides. *Acta Chem Scand* 4, 283–293.
- Edman, P., and Begg, G. (1967). A protein sequenator. *Eur J Biochem* 1, 80–91.
- El-Sharoud, W.M., and Rowbury, R.J. (2006). Recent insights into microbial physiology. *Sci Prog* 89, 141–149.
- Encheva, V., Gharbia, S.E., Wait, R., Begum, S., and Shah, H.N. (2006). Comparison of extraction procedures for proteome analysis of *Streptococcus pneumoniae* and a basic reference map. *Proteomics* 6, 3306–3317.
- Encheva, V., Wait, R., Gharbia, S.E., Begum, S., and Shah, H.N. (2005). Proteome analysis of serovars Typhimurium and Pullorum of *Salmonella enterica* subspecies I. *BMC Microbiol* 5, 42.
- Eng, J.K., McCormack, A.L., and Yates, J.R. (1994). An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J Am Soc Mass Spectrom* 5, 976–989.
- Eymann, C., Dreisbach, A., Albrecht, D., Bernhardt, J., Becher, D., Gentner, S., Tam, T., Büttner, K., Buurman, G., Scharf, C., et al. (2004). A comprehensive proteome map of growing *Bacillus subtilis* cells. *Proteomics* 4, 2849–2876.
- Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F., and Whitehouse, C.M. (1989). Electrospray ionization for mass spectrometry of large biomolecules. *Science* 246, 64–71.
- Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A., Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.F., Dougherty, B.A., Merrick, J.M., et al. (1995). Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269, 496–512.
- Folio, P., Chavant, P., Chafsey, I., Belkorchia, A., Chambon, C., and Hébraud, M. (2004). Two-dimensional electrophoresis database of *Listeria monocytogenes* EGD-e proteome and proteomic analysis of mid-log and stationary growth phase cells. *Proteomics* 4, 3187–3201.
- Fountoulakis, M., and Gasser, R. (2003). Proteomic analysis of the cell envelope fraction of *Escherichia coli*. *Amino Acids* 24, 19–41.
- Franco, A.T., Friedman, D.B., Nagy, T.A., Romero-Gallo, J., Krishna, U., Kendall, A., Israel, D.A., Tegtmeyer, N., Washington, M.K., and Peek, R.M. Jr. (2009). Delineation of a carcinogenic *Helicobacter pylori* proteome. *Mol Cell Proteomics* 8, 1947–1958.
- Gade, D., Gobom, J., and Rabus, R. (2005). Proteomic analysis of carbohydrate catabolism and regulation in the marine bacterium *Rhodospirillum rubrum*. *Proteomics* 5, 3672–3683.
- Gao, H., Yang, Z.K., Wu, L., Thompson, D.K., and Zhou, J. (2006). Global transcriptome analysis of the cold shock response of *Shewanella oneidensis* MR-1 and mutational analysis of its classical cold shock proteins. *J Bacteriol* 188, 4560–4569.
- Gardy, J.L., and Brinkman, F.S. (2006). Methods for predicting bacterial protein subcellular localization. *Nat Rev Microbiol* 4, 741–751.
- Garrels, J.I. (1989). The QUEST system for quantitative analysis of two-dimensional gels. *J Biol Chem* 264, 5269–5282.
- Gatlin, C.L., Pieper, R., Huang, S.T., Mongodin, E., Gebregeorgis, E., Parmar, P.P., Clark, D.J., Alami, H., Papazisi, L., Fleischmann, R.D., et al. (2006). Proteomic profiling of cell envelope-associated proteins from *Staphylococcus aureus*. *Proteomics* 6, 1530–1549.
- Gevaert, K., Van Damme, P., Ghesquière, B., Impens, F., Martens, L., Helsen, K., and Vandekerckhove, J. (2007). A la carte proteomics with an emphasis on gel-free techniques. *Proteomics* 7, 2698–2718.
- Görg, A., Obermaier, C., Boguth, G., Harder, A., Scheibe, B., Wildgruber, R., and Weiss, W. (2000). The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 21, 1037–1053.
- Greenough, C., Jenkins, R.E., Kitteringham, N.R., Pirmohamed, M., Park, B.K., and Pennington, S.R. (2004). A method for the rapid depletion of albumin and immunoglobulin from human plasma.

- Proteomics 4, 3107–3111.
- Guerrera, I.C., and Kleiner, O. (2005). Application of mass spectrometry in proteomics. *Biosci Rep* 25, 71–93.
- Gupta, M.K., Subramanian, V., and Yadav, J.S. (2009). Immunoproteomic identification of secretory and subcellular protein antigens and functional evaluation of the secretome fraction of *Mycobacterium immunogenum*, a newly recognized species of the *Mycobacterium chelonae-Mycobacterium abscessus* group. *J Proteome Res* 8, 2319–2330.
- Haas, G., Karaali, G., Ebermayer, K., Metzger, W.G., Lamer, S., Zimny-Arndt, U., Diescher, S., Goebel, U.B., Vogt, K., Roznowski, A.B., et al. (2002). Immunoproteomics of *Helicobacter pylori* infection and relation to gastric disease. *Proteomics* 2, 313–324.
- Han, M.J., and Lee, S.Y. (2006). The *Escherichia coli* proteome: past, present, and future prospects. *Microbiol Mol Biol Rev* 70, 362–439.
- Hecker, M., Antelmann, H., Büttner, K., and Bernhardt, J. (2008). Gel-based proteomics of Gram-positive bacteria: a powerful tool to address physiological questions. *Proteomics* 8, 4958–4975.
- Hecker, M., and Völker, U. (2004). Towards a comprehensive understanding of *Bacillus subtilis* cell physiology by physiological proteomics. *Proteomics* 4, 3727–3750.
- Henderson, B., and Martin, A. (2011). Bacterial virulence in the moonlight: multitasking bacterial moonlighting proteins are virulence determinants in infectious disease. *Infect Immun* 79, 3476–3491.
- Henzel, W.J., Watanabe, C., and Stults, J.T. (2003). Protein identification: the origins of peptide mass fingerprinting. *J Am Soc Mass Spectrom* 14, 931–942.
- Houthaeve, T., Gausepohl, H., Ashman, K., Nillson, T., and Mann, M. (1997). Automated protein preparation techniques using a digest robot. *J Protein Chem* 16, 343–348.
- Hueck, C.J. (1998). Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol Mol Biol Rev* 62, 379–433.
- Ishihama, Y., Schmidt, T., Rappsilber, J., Mann, M., Hartl, F.U., Kerner, M.J., and Frishman, D. (2008). Protein abundance profiling of the *Escherichia coli* cytosol. *BMC Genomics* 9, 102.
- James, P., Quadroni, M., Carafoli, E., and Gonnet, G. (1994). Protein identification in DNA databases by peptide mass fingerprinting. *Protein Sci* 3, 1347–1350.
- Jeffery, C.J. (2009). Moonlighting proteins—an update. *Mol Biosyst* 5, 345–350.
- Jongbloed, J.D., Martin, U., Antelmann, H., Hecker, M., Tjalsma, H., Venema, G., Bron, S., van Dijk, J.M., and Müller, J. (2000). TatC is a specificity determinant for protein secretion via the twin-arginine translocation pathway. *J Biol Chem* 275, 41350–41357.
- Jungblut, P.R. (2001). Proteome analysis of bacterial pathogens. *Microbes Infect* 3, 831–840.
- Jungblut, P.R., Bumann, D., Haas, G., Zimny-Arndt, U., Holland, P., Lamer, S., Siejak, F., Aebischer, A., and Meyer, T.F. (2000). Comparative proteome analysis of *Helicobacter pylori*. *Mol Microbiol* 36, 710–725.
- Kalia, A., and Gupta, R.P. (2005). Proteomics: a paradigm shift. *Crit Rev Biotechnol* 25, 173–198.
- Karas, M., and Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem* 60, 2299–2301.
- Klose, J. (1975). Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik* 26, 231–243.
- Klose, J. (2009). From 2-D electrophoresis to proteomics. *Electrophoresis* 30, S142–S149.
- Konkel, M.E., and Tilly, K. (2000). Temperature-regulated expression of bacterial virulence genes. *Microbes Infect* 2, 157–166.
- Lau, S.K., Fan, R.Y., Ho, T.C., Wong, G.K., Tsang, A.K., Teng, J.L., Chen, W., Watt, R.M., Curree, S.O., Tse, H., et al. (2011). Environmental adaptability and stress tolerance of *Laribacter hongkongensis*: a genome-wide analysis. *Cell Biosci* 1, 22.
- Lau, S.K., Woo, P.C., Fan, R.Y., Ma, S.S., Hui, W.T., Au, S.Y., Chan, L.L., Chan, J.Y., Lau, A.T., Leung, K.Y., et al. (2007). Isolation of *Laribacter hongkongensis*, a novel bacterium associated with gastroenteritis, from drinking water reservoirs in Hong Kong. *J Appl Microbiol* 103, 507–515.
- Lau, S.K.P., Lee, L.C.K., Fan, R.Y., Teng, J.L., Tse, C.W.S., Woo, P.C.Y., and Yuen, K.-Y. (2009). Isolation of *Laribacter hongkongensis*, a novel bacterium associated with gastroenteritis, from Chinese tiger frog. *Int J Food Microbiol* 129, 78–82.
- Lescuyer, P., Hochstrasser, D.F., and Sanchez, J.C. (2004). Comprehensive proteome analysis by chromatographic protein pre-fractionation. *Electrophoresis* 25, 1125–1135.
- Liao, X., Ying, T., Wang, H., Wang, J., Shi, Z., Feng, E., Wei, K., Wang, Y., Zhang, X., Huang, L., et al. (2003). A two-dimensional proteome map of *Shigella flexneri*. *Electrophoresis* 24, 2864–2882.
- Link, A.J., Eng, J., Schieltz, D.M., Carmack, E., Mize, G.J., Morris, D.R., Garvik, B.M., and Yates, J.R. 3rd. (1999). Direct analysis of protein complexes using mass spectrometry. *Nat Biotechnol* 17, 676–682.
- Lopez-Campistrous, A., Semchuk, P., Burke, L., Palmer-Stone, T., Broxk, S.J., Broderick, G., Bottorff, D., Bolch, S., Weiner, J.H., and Ellison, M.J. (2005). Localization, annotation, and comparison of the *Escherichia coli* K-12 proteome under two states of growth. *Mol Cell Proteomics* 4, 1205–1209.
- Macek, B., Mijakovic, I., Olsen, J.V., Gnad, F., Kumar, C., Jensen, P.R., and Mann, M. (2007). The serine/threonine/tyrosine phosphoproteome of the model bacterium *Bacillus subtilis*. *Mol Cell Proteomics* 6, 697–707.
- Marceau, M. (2005). Transcriptional regulation in *Yersinia*: an update. *Curr Issues Mol Biol* 7, 151–177.
- Markert, S., Arndt, C., Felbeck, H., Becher, D., Sievert, S.M., Hügler, M., Albrecht, D., Robidart, J., Bench, S., Feldman, R.A., et al. (2007). Physiological proteomics of the uncultured endosymbiont of *Riftia pachyptila*. *Science* 315, 247–250.
- Marouga, R., David, S., and Hawkins, E. (2005). The development of the DIGE system: 2D fluorescence difference gel analysis technology. *Anal Bioanal Chem* 382, 669–678.
- McHugh, L., and Arthur, J.W. (2008). Computational methods for protein identification from mass spectrometry data. *PLoS Comput Biol* 4, e12.
- Medberry, S., Gallagher, S., and Moomaw, B. (2005). Overview of digital electrophoresis analysis. *Curr Protoc Protein Sci Chapter* 10, Unit 10 12.
- Miesel, L., Greene, J., and Black, T.A. (2003). Genetic strategies for antibacterial drug discovery. *Nat Rev Genet* 4, 442–456.

- Molloy, M.P., Herbert, B.R., Slade, M.B., Rabilloud, T., Nouwens, A.S., Williams, K.L., and Gooley, A.A. (2000). Proteomic analysis of the *Escherichia coli* outer membrane. *Eur J Biochem* 267, 2871–2881.
- Mora, M., Donati, C., Medini, D., Covacci, A., and Rappuoli, R. (2006). Microbial genomes and vaccine design: refinements to the classical reverse vaccinology approach. *Curr Opin Microbiol* 9, 532–536.
- Morris, H.R., Panico, M., Barber, M., Bordoli, R.S., Sedgwick, R.D., and Tyler, A. (1981). Fast atom bombardment: a new mass spectrometric method for peptide sequence analysis. *Biochem Biophys Res Commun* 101, 623–631.
- Ni, X.P., Ren, S.H., Sun, J.R., Xiang, H.Q., Gao, Y., Kong, Q.X., Cha, J., Pan, J.C., Yu, H., and Li, H.M. (2007). *Laribacter hongkongensis* isolated from a patient with community-acquired gastroenteritis in Hangzhou City. *J Clin Microbiol* 45, 255–256.
- Nouwens, A.S., Willcox, M.D., Walsh, B.J., and Cordwell, S.J. (2002). Proteomic comparison of membrane and extracellular proteins from invasive (PAO1) and cytotoxic (6206) strains of *Pseudomonas aeruginosa*. *Proteomics* 2, 1325–1346.
- O'Connor, C.D., Farris, M., Fowler, R., and Qi, S.Y. (1997). The proteome of *Salmonella enterica* serovar *typhimurium*: current progress on its determination and some applications. *Electrophoresis* 18, 1483–1490.
- O'Farrell, P.H. (1975). High resolution two-dimensional electrophoresis of proteins. *J Biol Chem* 250, 4007–4021.
- Pancholi, V., and Fischetti, V.A. (1992). A major surface protein on group A *streptococci* is a glyceraldehyde-3-phosphate-dehydrogenase with multiple binding activity. *J Exp Med* 176, 415–426.
- Parkhill, J., and Wren, B.W. (2011). Bacterial epidemiology and biology - lessons from genome sequencing. *Genome Biol* 12, 230.
- Patton, W.F. (2002). Detection technologies in proteome analysis. *J Chromatogr B Analyt Technol Biomed Life Sci* 771, 3–31.
- Pemberton, J.M., Kidd, S.P., and Schmidt, R. (1997). Secreted enzymes of *Aeromonas*. *FEMS Microbiol Lett* 152, 1–10.
- Perkins, D.N., Pappin, D.J.C., Creasy, D.M., and Cottrell, J.S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20, 3551–3567.
- Perrin, C., González-Márquez, H., Gaillard, J.L., Bracquart, P., and Guimont, C. (2000). Reference map of soluble proteins from *Streptococcus thermophilus* by two-dimensional electrophoresis. *Electrophoresis* 21, 949–955.
- Phadtare, S., Alsina, J., and Inouye, M. (1999). Cold-shock response and cold-shock proteins. *Curr Opin Microbiol* 2, 175–180.
- Phillips, C.I., and Bogoy, M. (2005). Proteomics meets microbiology: technical advances in the global mapping of protein expression and function. *Cell Microbiol* 7, 1061–1076.
- Poetsch, A., and Wolters, D. (2008). Bacterial membrane proteomics. *Proteomics* 8, 4100–4122.
- Poland, T., Rabilloud, T., and Sinha, P. (2005). Silver Staining of 2-D Gels. In: *The Proteomics Protocols Handbook*, Walker J.M., ed. Totowa, NJ: Human Press, 177–184.
- Qian, W.J., Jacobs, J.M., Liu, T., Camp, D.G. 2nd, and Smith, R.D. (2006). Advances and challenges in liquid chromatography-mass spectrometry-based proteomics profiling for clinical applications. *Mol Cell Proteomics* 5, 1727–1744.
- Rabilloud, T. (2009). Membrane proteins and proteomics: love is possible, but so difficult. *Electrophoresis* 30, S174–S180.
- Rabilloud, T., Chevallet, M., Lucche, S., and Lelong, C. (2010). Two-dimensional gel electrophoresis in proteomics: Past, present and future. *J Proteomics* 73, 2064–2077.
- Rabilloud, T., Heller, M., Gasnier, F., Lucche, S., Rey, C., Aebbersold, R., Benahmed, M., Louisot, P., and Lunardi, J. (2002). Proteomics analysis of cellular response to oxidative stress. Evidence for *in vivo* overoxidation of peroxiredoxins at their active site. *J Biol Chem* 277, 19396–19401.
- Rabilloud, T., Vaezzadeh, A.R., Potier, N., Lelong, C., Leize-Wagner, E., and Chevallet, M. (2009). Power and limitations of electrophoretic separations in proteomics strategies. *Mass Spectrom Rev* 28, 816–843.
- Rappuoli, R. (2001). Reverse vaccinology, a genome-based approach to vaccine development. *Vaccine* 19, 2688–2691.
- Regula, J.T., Ueberle, B., Boguth, G., Görg, A., Schnölzer, M., Herrmann, R., and Frank, R. (2000). Towards a two-dimensional proteome map of *Mycoplasma pneumoniae*. *Electrophoresis* 21, 3765–3780.
- Renzone, G., D'Ambrosio, C., Arena, S., Rullo, R., Ledda, L., Ferrara, L., and Scaloni, A. (2005). Differential proteomic analysis in the study of prokaryotes stress resistance. *Ann Ist Super Sanita* 41, 459–468.
- Roe, M.R., and Griffin, T.J. (2006). Gel-free mass spectrometry-based high throughput proteomics: tools for studying biological response of proteins and proteomes. *Proteomics* 6, 4678–4687.
- Rosen, R., Sacher, A., Shechter, N., Becher, D., Büttner, K., Biran, D., Hecker, M., and Ron, E.Z. (2004). Two-dimensional reference map of *Agrobacterium tumefaciens* proteins. *Proteomics* 4, 1061–1073.
- Rosengren, A.T., Salmi, J.M., Aittokallio, T., Westerholm, J., Lahesmaa, R., Nyman, T.A., and Nevalainen, O.S. (2003). Comparison of PDQuest and Progenesis software packages in the analysis of two-dimensional electrophoresis gels. *Proteomics* 3, 1936–1946.
- Sabarth, N., Hurwitz, R., Meyer, T.F., and Bumann, D. (2002). Multi-parameter selection of *Helicobacter pylori* antigens identifies two novel antigens with high protective efficacy. *Infect Immun* 70, 6499–6503.
- Santoni, V., Molloy, M., and Rabilloud, T. (2000). Membrane proteins and proteomics: un amour impossible? *Electrophoresis* 21, 1054–1070.
- Sarioglu, H., Lottspeich, F., Walk, T., Jung, G., and Eckerskorn, C. (2000). Deamidation as a widespread phenomenon in two-dimensional polyacrylamide gel electrophoresis of human blood plasma proteins. *Electrophoresis* 21, 2209–2218.
- Scheele, G.A. (1975). Two-dimensional gel analysis of soluble proteins. Characterization of guinea pig exocrine pancreatic proteins. *J Biol Chem* 250, 5375–5385.
- Sibbald, M.J., Ziebandt, A.K., Engelmann, S., Hecker, M., de Jong, A., Harmsen, H.J., Raangs, G.C., Stokroos, I., Arends, J.P., Dubois, J.Y., et al. (2006). Mapping the pathways to staphylococcal pathogenesis by comparative secretomics. *Microbiol Mol Biol Rev* 70, 755–788.
- Skurnik, M., Venho, R., Bengoechea, J.A., and Moriyón, I. (1999). The lipopolysaccharide outer core of *Yersinia enterocolitica* serotype O:3 is required for virulence and plays a role in outer membrane integrity. *Mol Microbiol* 31, 1443–1462.
- Swanson, R.V., Alex, L.A., and Simon, M.I. (1994). Histidine and

- aspartate phosphorylation: two-component systems and the limits of homology. *Trends Biochem Sci* 19, 485–490.
- Teng, J.L., Woo, P.C., Ma, S.S., Sit, T.H., Ng, L.T., Hui, W.T., Lau, S.K., and Yuen, K.Y. (2005). Ecoepidemiology of *Laribacter hongkongensis*, a novel bacterium associated with gastroenteritis. *J Clin Microbiol* 43, 919–922.
- Thein, M., Sauer, G., Paramasivam, N., Grin, I., and Linke, D. (2010). Efficient subfractionation of gram-negative bacteria for proteomics studies. *J Proteome Res* 9, 6135–6147.
- Thieringer, H.A., Jones, P.G., and Inouye, M. (1998). Cold shock and adaptation. *Bioessays* 20, 49–57.
- Tjalsma, H., Antelmann, H., Jongbloed, J.D., Braun, P.G., Darmon, E., Dorenbos, R., Dubois, J.Y., Westers, H., Zanen, G., Quax, W.J., et al. (2004). Proteomics of protein secretion by *Bacillus subtilis*: separating the “secrets” of the secretome. *Microbiol Mol Biol Rev* 68, 207–233.
- Traini, M., Gooley, A.A., Ou, K., Wilkins, M.R., Tonella, L., Sanchez, J.C., Hochstrasser, D.F., and Williams, K.L. (1998). Towards an automated approach for protein identification in proteome projects. *Electrophoresis* 19, 1941–1949.
- Trost, M., Wehmhöner, D., Kärst, U., Dieterich, G., Wehland, J., and Jänsch, L. (2005). Comparative proteome analysis of secretory proteins from pathogenic and nonpathogenic *Listeria* species. *Proteomics* 5, 1544–1557.
- Trülzsch, K., Roggenkamp, A., Aepfelbacher, M., Wilharm, G., Ruckdeschel, K., and Heesemann, J. (2003). Analysis of chaperone-dependent Yop secretion/translocation and effector function using a mini-virulence plasmid of *Yersinia enterocolitica*. *Int J Med Microbiol* 293, 167–177.
- Unlü, M., Morgan, M.E., and Minden, J.S. (1997). Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* 18, 2071–2077.
- Völker, U., and Hecker, M. (2005). From genomics via proteomics to cellular physiology of the Gram-positive model organism *Bacillus subtilis*. *Cell Microbiol* 7, 1077–1085.
- Wang, J., Ying, T., Wang, H., Shi, Z., Li, M., He, K., Feng, E., Wang, J., Yuan, J., Li, T., et al. (2005). 2-D reference map of *Bacillus anthracis* vaccine strain A16R proteins. *Proteomics* 5, 4488–4495.
- Wang, Y., Xu, A., Knight, C., Xu, L.Y., and Cooper, G.J. (2002). Hydroxylation and glycosylation of the four conserved lysine residues in the collagenous domain of adiponectin. Potential role in the modulation of its insulin-sensitizing activity. *J Biol Chem* 277, 19521–19529.
- Washburn, M.P., Wolters, D., and Yates, J.R. 3rd. (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 19, 242–247.
- Wilkins, M.R., Gasteiger, E., Sanchez, J.C., Bairoch, A., and Hochstrasser, D.F. (1998). Two-dimensional gel electrophoresis for proteome projects: the effects of protein hydrophobicity and copy number. *Electrophoresis* 19, 1501–1505.
- Wilkins, M.R., Pasquali, C., Appel, R.D., Ou, K., Golaz, O., Sanchez, J.C., Yan, J.X., Gooley, A.A., Hughes, G., Humphery-Smith, I., et al. (1996). From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Biotechnology (N Y)* 14, 61–65.
- Woo, P.C., Lau, S.K., Teng, J.L., Que, T.L., Yung, R.W., Luk, W.K., Lai, R.W., Hui, W.T., Wong, S.S., Yau, H.H., et al., and the L Hongkongensis study group. (2004). Association of *Laribacter hongkongensis* in community-acquired gastroenteritis with travel and eating fish: a multicentre case-control study. *Lancet* 363, 1941–1947.
- Woo, P.C., Lau, S.K., Teng, J.L., and Yuen, K.Y. (2005). Current status and future directions for *Laribacter hongkongensis*, a novel bacterium associated with gastroenteritis and traveller’s diarrhoea. *Curr Opin Infect Dis* 18, 413–419.
- Woo, P.C., Lau, S.K., Tse, H., Teng, J.L., Curreem, S.O., Tsang, A.K., Fan, R.Y., Wong, G.K., Huang, Y., Loman, N.J., et al. (2009). The complete genome and proteome of *Laribacter hongkongensis* reveal potential mechanisms for adaptations to different temperatures and habitats. *PLoS Genet* 5, e1000416.
- Yan, J.X., Devenish, A.T., Wait, R., Stone, T., Lewis, S., and Fowler, S. (2002). Fluorescence two-dimensional difference gel electrophoresis and mass spectrometry based proteomic analysis of *Escherichia coli*. *Proteomics* 2, 1682–1698.
- Yates, J.R. 3rd, Speicher, S., Griffin, P.R., and Hunkapiller, T. (1993). Peptide mass maps: a highly informative approach to protein identification. *Anal Biochem* 214, 397–408.
- Yu, H.B., Kaur, R., Lim, S., Wang, X.H., and Leung, K.Y. (2007). Characterization of extracellular proteins produced by *Aeromonas hydrophila* AH-1. *Proteomics* 7, 436–449.
- Yuen, K.Y., Woo, P.C., Teng, J.L., Leung, K.W., Wong, M.K., and Lau, S.K. (2001). *Laribacter hongkongensis* gen. nov., sp. nov., a novel gram-negative bacterium isolated from a cirrhotic patient with bacteremia and empyema. *J Clin Microbiol* 39, 4227–4232.