Molecular Phylogenetics and Evolution 127 (2018) 239–247

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Global evolution of glycosylated polyene macrolide antibiotic biosynthesis

Dongmei Liang[1], Xiaoyu Yang[1], Jiaheng Liu, Qinggele Caiyin, Guangrong Zhao, Lifen Li, Jianjun Qiao*

Department of Pharmaceutical Engineering, School of Chemical Engineering and Technology, Tianjin University, Tianjin 300072, PR China
Key Laboratory of Systems Bioengineering (Ministry of Education), Tianjin University, Tianjin 300072, PR China
SynBio Research Platform, Collaborative Innovation Center of Chemical Science and Engineering (Tianjin), Tianjin 300072, PR China

ARTICLE INFO

Keywords:
Glycosylated polyene macrolides
Global evolution
Antibiotic biosynthesis
Actinomycetes genomes
Host preference

ABSTRACT

Antibiotics are the most marvelous evolutionary products of microbes to obtain competitive advantage and maintain ecological balance. However, the origination and development of antibiotics has yet to be explicitly investigated. Due to diverse structures and similar biosynthesis, glycosylated polyene macrolides (gPEMs) were chosen to explore antibiotic evolution. A total of 130 candidate and 38 transitional gPEM clusters were collected from actinomycetes genomes, providing abundant references for phenotypic gaps in gPEM evolution. The most conserved parts of gPEM biosynthesis were found and used for phylogeny construction. On this basis, we proposed ancestral gPEM clusters at different evolutionary stages and interpreted the possible evolutionary histories in detail. The results revealed that gPEMs evolved from small rings to large rings and continuously increased structural diversity through acquiring, discarding and exchanging genes from different evolutionary origins, as well as co-evolution of functionally related proteins. The combination of horizontal gene transfers, environmental effects and host preference resulted in the diversity and worldwide distribution of gPEMs. This study is not only a useful exploration on antibiotic evolution but also an inspiration for diversity and biogeographic investigations on antibiotics in the era of Big Data.

1. Introduction

Antibiotics are the most important secondary metabolites of environmental microbes. They can not only eliminate or inhibit the competitors in living environments, but also play signaling and regulatory roles in microbial communities (Aminov, 2009). Meanwhile, many types of antibiotics are essential for ecological balance and harmonious coexistence of the producers and their hosts. Polyene macrolides (PEMs) are a series of widespread polyketide antibiotics (Jørgensen et al., 2009) as the main anti-pathogenic fungal agents of many actinomycetes in diverse environments, as well as the primary weapons (from symbiotic actinomycetes) used by leaf-cutting ants to protect their fungus culture gardens (Haeder et al., 2009). Moreover, some of PEMs have been successfully developed as antifungal drugs and food preservatives (Caffrey et al., 2016).

PEMs possess large macrolactone rings with a series of conjugated double bonds and can be classified according to the modification patterns: the glycosylated PEMs (gPEMs) such as nystatin (Fig. 1b) and the non-glycosylated PEMs (nPEMs) such as filipin. Currently, most of PEMs isolated from nature and widely used in clinics and industries are gPEMs, which have been well studied in biosynthetic pathways. To date, a total of ten gPEM biosynthetic gene clusters have been completely sequenced, designated as *nys* (for nystatin) (Fjaervik & Zotchev, 2005), *amph* (for amphotericin) (Caffrey et al., 2001), *cpp* (for NPP) (Kim et al., 2009), *fsc* (for FR-008/candicidin) (Chen et al., 2003), *pim* (Aparicio et al., 2000)/sgn (Wang et al., 2016)/*scn* (Du et al., 2011) (for pimaricin), *tetr* (Cao et al., 2012)/*ttm* (Ren et al., 2014) (for tetramycin) and *ace* (for 67-121C) (Sheehan et al., 2017). Also, *per* (for perimycin) (Hutchinson et al., 2010), *rim* (for rimocidin/CE-108) (Seco et al., 2004), *can* (for candicidin) (Gil & Campelo-Diez, 2003) and *nyp* (for nystatin P1) (Barke et al., 2010) were partially identified (Fig. 1a). We have noticed that almost all of the natural gPEMs used as antibiotics or preservatives are 26/28-membered rings with tetraene structures or 38-membered rings with heptaene/heptaene-like structures. Interestingly, although there are large differences among the macrolactone rings, the basic skeleton and tailoring parts are quite conserved, containing an exocyclic carboxyl group and an unusual sugar moiety named mycosamine (Fig. 1b). As expected, all identified gPEM biosynthetic pathways are similar substantially.

gPEM biosynthesis begins with a repetitive decarboxylative

* Corresponding author. School of Chemical Engineering and Technology, Tianjin University, Tianjin 300072, PR China.
E-mail address: jianjunq@tju.edu.cn (J. Qiao).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ympev.2018.06.005
Received 12 February 2018; Accepted 4 June 2018
Available online 07 June 2018
1055-7903/

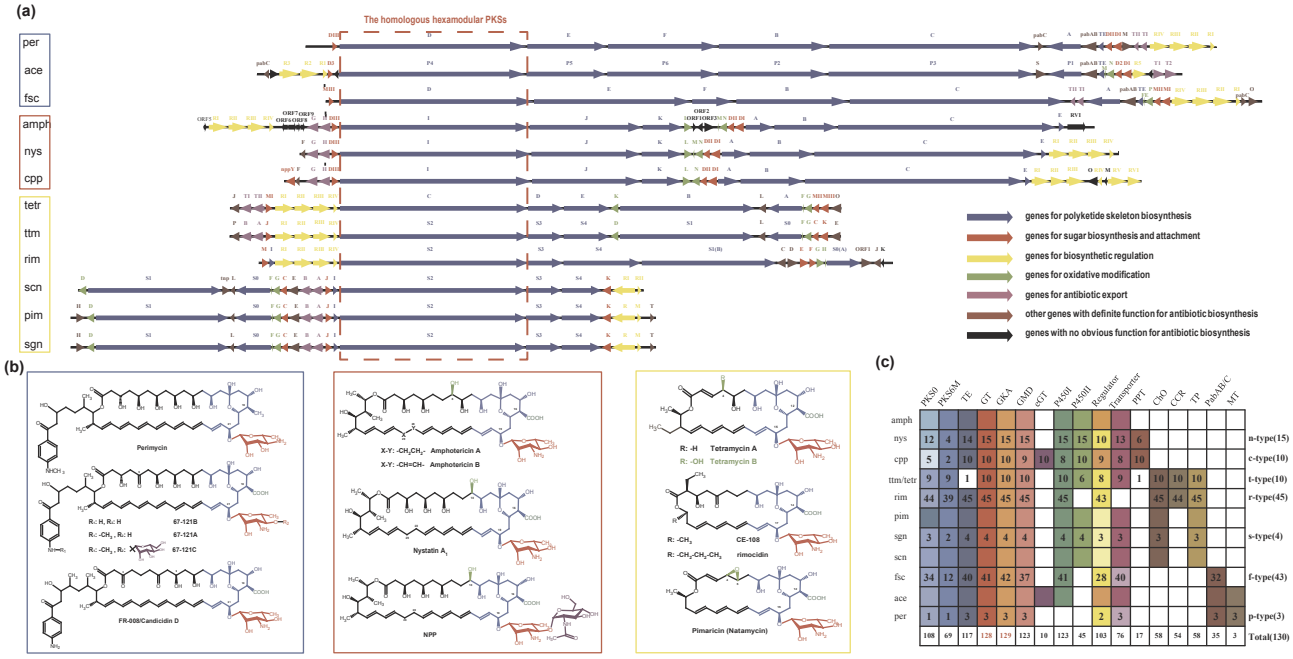

Fig. 1. The overview of identified gPEMs. (a) gPEM biosynthetic gene clusters aligned by the common hexamodular PKSs. (b) Structures of representative gPEMs. The partial skeletons highlighted in blue are biosynthesized by the homologous hexamodular PKSs. The tailoring parts marked in dark green, light green, red and purple are formed by P450Is, P450IIs, GTs and eGTs correspondingly. (c) Conservation statistics of gPEM biosynthetic proteins in identified and candidate gPEM clusters. Different colors represent different protein forms in identified gPEM biosynthesis (according to the cluster names at the left), and the numbers represent intact ORFs conserved in candidate clusters (according to the cluster classifications at the right). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

condensation of an activated starter unit with methylmalonyl- or malonyl-CoA-derived extender units to form a growing polyketide chain, catalyzed by type I polyketide synthases (PKSs). Following the cyclization of polyketide chain, a type of cytochrome P450 (P450I) participates in the formation of the carboxyl group. GDP-mannose 4, 6 dehydratase (GMD) and GDP-ketosugar aminotransferase (GKA) take part in sugar biosynthesis and the sugar moiety is attached to the macrolactone ring by glycosyltransferase (GT) after the formation of the carboxyl group. Then oxidation in the polyol region (if necessary) is completed by another type of P450 (P450II). The similar biosynthetic processes and widespread distribution provide more possibilities to explore how gPEMs got so variform structures and spread to every corner of the world.

To investigate the development of gPEM biosynthetic pathways in nature, we searched for the possible gPEM clusters in current actinomycetes genome database, collecting considerable candidate and transitional clusters, providing abundant phenotypic transitions for gPEM evolution. Based on the comparative analysis of gPEM clusters, we discovered the most conserved parts of gPEM biosynthesis and constructed phylogeny using best-preserved GT and GKA. Accordingly, ancestral gPEM clusters at different evolutionary stages were proposed. Then possible evolutionary histories of gPEMs were also discussed in detail. The results revealed that gPEMs evolved from small rings to large rings and continuously increased structural diversity for better bioactivities or host preference. Meanwhile, co-evolution of functionally related proteins and environmental effects balanced the conservation and variation during evolutionary process. Moreover, frequent horizontal gene transfers (HGTs) might play fundamental roles in distribution of gPEMs. Our studies not only interpreted the origination and development of gPEMs but also provided an efficient alternative method for investigating the evolutionary processes of specific metabolites through mining current massive data.

2. Material and methods

2.1. Genomic sequences and resources

All identified gPEM clusters were retrieved from GenBank (Table S1) and sorted out according to relevant literatures. Sequences of gPEM biosynthetic proteins were obtained from the National Center for Biotechnology Information (NCBI: http://www.ncbi.nlm.nih.gov/) (Table S2).

2.2. Identification of candidate and transitional gPEM clusters

Putative gPEM cluster searches were done by using MultiGeneBlast (Medema et al., 2013). Chromosome segments containing more than three types of homologous gPEM proteins (identities > 40%) were considered as potential clusters. Those largely resembled identified clusters were designated as “candidate clusters”. Then, they were classified according to the sequence identities and the cluster organizations and marked as n-(*nys*), c-(*cpp*), f-(*fsc*), r-(*rim*), s-(*pim/scn/sgn*), t-(*ttm/tetr*) and p-(*per*) type clusters respectively (Fig. S1 & Table S1). Those with almost entire cluster organization but resembled more than one identified clusters simultaneously were designated as “transitional (tr-) clusters”. Meanwhile, those with most of similar biosynthetic genes but do not cluster or the organizations obviously differ from any identified clusters were designated as “fragmental (fr-) clusters”. Information about environmental and other factors associated with strains was acquired from either NCBI or Genomes OnLine Database (GOLD) (Reddy et al., 2015).

2.3. Multiple sequence alignment and phylogeny construction

Standard approaches were used to reconstruct phylogenetic trees of KSs, GTs, GKAs and 16s rRNAs. GTs/GKAs used for phylogenetic analysis were listed in Table S3. The selected sequences were aligned by Clustal X 2.0 (Larkin et al., 2007) with default parameters, and the