

<<小分子RNA介导的基因表达调控>>

图书基本信息

书名：<<小分子RNA介导的基因表达调控>>

13位ISBN编号：9787030329097

10位ISBN编号：7030329090

出版时间：2012-1

出版时间：科学出版社

作者：Rajesh K.Gaur , Ph.D、 John J.Rossi , Ph.D 著，郭惠珊，段成国 译

页数：432

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

<<小分子RNA介导的基因表达调控>>

内容概要

内源小分子RNA广泛存在于各种生物中，包括人类、小鼠、果蝇、蠕虫、真菌和细菌等。microRNA作为一种细胞调控关键因子能够修饰基因的表达。在高等真核生物中，microRNA甚至能调控约50%基因的表达。

本书汇集了众多科技工作者的前沿性工作，内容包括从细菌到人类等生物组织中microRNA调控途径的多样性。除了阐述调控小分子RNA的生物合成机制及其加工过程，作者还探讨了这些途径的功能在寄主体内的重要性。

本书围绕小分子RNA这一新发现的调控分子，针对其参与调控的广度与创新性进行了阐述。小分子RNA已经成为研究基因功能的强有力工具，并带来了一系列的重大发现，必将对增进基因功能与疾病治疗的理解带来革命性的改变。

<<小分子RNA介导的基因表达调控>>

书籍目录

- 前言
致谢
编者简介
撰稿人
- 第1章 MicroMining:通过计算方式发现未知的microRNA Adam Grundhoff
第2章 动物microRNA基因预测 Ola Sn ̄ve,Pal S*tr̄om
第3章 研究microRNA存在与功能的一系列资源 Praveen Sethupathy,Molly Megraw, Artemis G. Hatzigeorgiou
第4章 大肠杆菌Hfq结合小RNA对mRNA稳定性及翻译的调控 Hiroji Aiba
第5章 动物细胞中microRNA调控基因表达的机制 Yang Yu,Timothy W. Nilsen
第6章 秀丽隐杆线虫microRNA Mona J. Nolde, Frank J. Slack
第7章 秀丽隐杆线虫小RNA的分离及鉴定 Chisato Ushida, Yusuke Hokii
第8章 MicroRNA与果蝇发育 Utpal Bhadra,Sunit KumarSingh,Singh,S. N. C. V. L. Pushpavalli,Praveensingh B. Hajeri,Manika Pal-Bhadra
第9章 斑马鱼RNA干扰与microRNA Alex S. Flynt,Elizabeth J. Thatcher,James G. Patton
第10章 植物microRNA的产生和功能 Zoltan Havelda
第11章 拟南芥内源小RNA途径 Manu Agarwal,Julien Curaba,Xuemei Chen
第12章 如何评价microRNA表达——技术指导 Mirco Castoldi,Vladimir Benes,Martina U. Muckenthaler
第13章 MicroRNA基因表达定量的方法 Lori A. Neely
第14章 MicroRNA介导的可变剪切调控 Rajesh K. Gaur
第15章 RNA聚合酶 介导的内含子microRNA表达系统研究进展 Shi-Lung Lin,Shao-Yao Ying
第16章 基于microRNA的RNA聚合酶 表达载体在动物细胞RNA干扰中的应用 Anne B. Vojtek,Kwan-Ho Chung,Paresh D. Patel,David L. Turner
第17章 转基因RNA干扰技术——一种用于哺乳动物反向遗传学研究的快速低成本方法 Linghua Qiu,Zuoshang Xu
第18章 AIDS交响曲——基于microRNA的治疗方法 Yoichi R. Fujii
第19章 MicroRNA与癌症——连点成线 Sumedha D. Jayasena
第20章 哺乳动物中microRNA介导的转录水平基因沉默 Daniel H. Kim, John J. Rossi
第21章 由RNA介导的转录水平基因沉默控制的基因表达调控 Kevin V. Morris
- 索引

<<小分子RNA介导的基因表达调控>>

章节摘录

1 MicroMining Computational Approaches to microRNA Discovery Adam Grundhoff Overview.....11.1
 Introduction.....21.2 When Is a Small RNA an miRNA?.....21.3 Advantages and Disadvantages of Experimental versus Computational miRNA Identification.....31.4 Computational Prediction of miRNAs.....51.4.1 Getting Started: Upstream Filtering.....71.4.2 Following Through: Structure Prediction and Scoring.....121.4.3 Wrapping It Up: Experimental Validation.....141.5 Viral miRNAs.....151.6
 Conclusions.....
 16 References..... 16 Overview

The recent past has seen the rapid identification of thousands of microRNAs (miRNAs) encoded by various metazoan organisms as well as some viruses, and it is very likely that many more still await discovery. Most of the hitherto-known miRNAs have been identified via the cloning and sequencing of small RNAs. While very powerful, this approach is not without its limitations: especially those miRNAs that are of low abundance, or which are only expressed in certain cell types or only during brief periods of organismal development, or are easily missed in cloning-based screens. Thus, alternative means of miRNA discovery are needed. Given that the signal that marks the miRNA precursor for the cellular processing machinery appears to be a relatively simple one (i.e., a hairpin structure), and considering the rapidly increasing availability of large-scale genomic sequencing data for many organisms, computational methods appear ideally suited for the comprehensive identification of hitherto-unknown miRNAs. This chapter discusses the general principles of computational miRNA identification methods, examines their advantages and disadvantages as compared to the cloning method, and takes a look at the various miRNA prediction algorithms that have been developed recently.

1.1 Introduction miRNAs are small (~ 22 nt) RNA molecules that are able to regulate the expression of fully or partially complementary mRNA transcripts. As described in greater detail elsewhere in this book, they are initially transcribed as part of hairpin structures within much larger precursor transcripts (the so-called primary RNAs or pri-miRNAs). Following excision of the stem-loops by the RNase III-like enzyme Drosha, the isolated hairpins (called precursor miRNAs or pre-miRNAs) are exported to the cytoplasm and further processed by the Dicer complex to produce the mature, single-stranded miRNA molecule. Recent evidence suggests that plants and animals encode a multitude of miRNAs, many of which are evolutionarily conserved. As of this writing, it is still true that the majority of known miRNAs have been identified experimentally, that is, by cloning of small RNAs. However, this method has certain limitations, and alternative means for the prediction of novel miRNAs are therefore increasingly sought. The observation that pre-miRNAs form characteristic stem-loops has spurred the development of a number of computational approaches designed to identify novel miRNA candidates based on the prediction and analysis of secondary structures. Given the already complete or near-complete sequencing of whole genomes from many species, such approaches hold great promise for identifying the full complement of miRNAs encoded by a given organism. However, because the precise set of structural features that differentiate a pre-miRNA stem-loop from the large number of hairpins in the genome is not known, additional filters have to be employed to reduce the number of false-positive predictions, and experimental confirmation of the remaining candidates is required. In this chapter, I will compare the benefits and disadvantages of computational miRNA prediction methods in comparison to the cloning method, review principles of the existing miRNA prediction algorithms, discuss the general challenges and pitfalls of in silico miRNA identification, and provide an outlook of what might be expected from these approaches in the future. Finally, I will consider a special application of the miRNA prediction problem: the identification of miRNAs in viral genomes.

1.2 When is a small RNA an miRNA? In order to devise approaches designed to identify miRNAs, be they experimental or computational, it is important to clearly define what an miRNA is. In a

<<小分子RNA介导的基因表达调控>>

biological sense, such a definition is quite straightforward: an miRNA is simply a small, single-stranded regulatory RNA molecule that is generated from its precursor molecules via successive processing by Drosha and Dicer. It is much more difficult, however, to define practicable criteria that are readily testable on an experimental or computational basis and that can unequivocally identify a candidate sequence as a genuine miRNA. Following the realization that miRNAs represent abundant molecules expressed in a wide variety of organisms, a consortium of researchers agreed on a set of criteria that have to be fulfilled before a candidate can be called a bona fide miRNA.¹ According to these guidelines, it is necessary to provide evidence that (1) the candidate sequence is expressed as an appropriately sized RNA molecule in living cells and, furthermore, does not stem from random degradation (Expression criteria), and (2) that the maturation of the candidate involves processing by Drosha and Dicer (Biogenesis criteria). The expression criteria are preferentially satisfied by detection of a distinct band of approximately 22 nt on a Northern blot. Alternatively, the ability to detect the molecule in a library of cloned, size-selected RNAs is considered sufficient evidence, especially if the library contains high copy numbers of the particular candidate sequence. To satisfy the biogenesis criteria, the guidelines by Ambros et al.¹ call for experimental proof of Dicer processing by demonstrating that increased levels of the precursor accumulate in cells with decreased Dicer expression. In contrast, experimental proof of Drosha processing is generally not required; instead, it is sufficient to show that the putative precursor transcript has the capacity to adopt a secondary structure that is likely to be amenable to Drosha processing. Of course, given the incomplete knowledge of the rules governing recognition of target mRNAs by Drosha, it is not known what exactly makes a given RNA structure amenable to Drosha processing, and (as will be discussed later) this complicates the computational prediction of miRNA candidates considerably. Based on the characteristics of known miRNA precursor structures, however, it is generally agreed that the minimal requirements are (1) the adopted structure is a hairpin that does not contain many or large internal bulges, and (2) the mature miRNA is to be found within the stem (not the loop) part of the hairpin. Evolutionary conservation serves as a third biogenesis criterion: As miRNAs are often conserved in closely related (and sometimes even in distant) species, phylogenetic conservation of the miRNA sequence itself as well as its fold-back structure is considered strong evidence that the candidate sequence represents a genuine miRNA. An ideal miRNA candidate would meet all of the preceding criteria; however, it is generally considered sufficient to provide convincing evidence for at least one criterion out of the two categories. Indeed, because Dicer knockout cells are not readily available for most organisms, and effective knockdown of Dicer is technically challenging, positive experimental proof of Dicer processing is rarely shown.

1.3 Advantages and Disadvantages of Experimental versus Computational miRNA Identification

The “ traditional ” approach to identifying miRNAs consists of cloning of small RNA moieties. Although several protocols for the efficient cloning of such molecules have been devised, they all rely on the common principle of ligating linkers to size-fractionated RNAs, followed by cDNA synthesis and typically PCR amplification. The obtained products are then either cloned (often after concatamerization to increase the information obtained in a single-sequence read) and sequenced, or subjected directly to massive parallel sequencing approaches (“ deep sequencing ”). According to the guidelines described earlier, these candidates are then further evaluated to ensure that the putative pre-miRNA sequence adopts an appropriate hairpin structure around the candidate. If this is the case, the candidate can generally be considered a bona fide miRNA, since the recovery of the clone from a small RNA library already satisfies the expression criterion (nevertheless, Northern blots are often performed to allow for proper quantification of the miRNA). The cloning approach has been extremely successful, and although increasing numbers of miRNAs are being identified via computational means, the majority of confirmed miRNAs currently listed in the miRNA database (miRBase, <http://microrna.sanger.ac.uk>) still have been identified via this method. One of the great advantages of the cloning protocol is that it provides the precise sequence of the mature miRNA molecule. Therefore, in contrast to hybridization-based methods, even closely related miRNAs that differ in only one nucleotide position can be distinguished. Also, the currently available computational prediction tools generally only allow identification of miRNA precursors but do not reliably predict the location of Drosha and Dicer cleavage sites. In contrast, cloning identifies the precise 5' and 3' termini of the mature miRNA molecule. As it appears

<<小分子RNA介导的基因表达调控>>

that nucleotides 2 to 8 of the miRNA (the so-called seed region) are especially important for target recognition, knowledge of the precise ends (and particularly the 5' terminus) is a distinct advantage if a computational prediction of target transcripts is to be performed. As might be expected, the frequency with which a given miRNA is cloned often is approximately equivalent to its abundance (although this frequency may also be affected by other factors; see the following text) and therefore provides a rough estimate of its expression levels. Thus, abundantly expressed miRNAs are usually readily identified. However, it can be challenging to achieve a saturated screen that also captures rare miRNAs. Furthermore, even if such miRNAs are contained within the library, one can never be entirely certain that enough clones have been sequenced to identify all of them. In addition to these constraints, the scope of a cloning screen is also limited by its source material; naturally, only miRNAs that are expressed in the cells from which the RNA material was derived can be identified. Many miRNAs, however, are expressed in a tissue-dependent manner, or are only expressed at certain developmental stages. This limitation can be partially overcome for relatively simple organisms, where the RNA can be prepared from whole animals (e.g., mixed larvae stages and adults from worms or insects). In organisms with higher complexity such as vertebrates, however, the situation is more difficult: RNA from different embryonic or adult tissues can be mixed, but the sensitivity of the screen will dramatically decrease with the complexity of the source material, and it is very unlikely that nonabundant miRNAs could be identified in such screens. While these problems could be theoretically solved by massive screening efforts, that is, performing separate screens with material prepared from every individual tissue at each developmental stage, the cloning approach also appears limited in a more fundamental way. Several observations suggest that some miRNAs are more readily cloned than others owing to intrinsic properties such as sequence composition, the presence of certain nucleotides at their termini, or posttranscriptional modifications such as methylation or RNA editing.²⁶ Computational approaches to miRNA discovery are not subject to many of the limitations that apply to the cloning method. Certainly, one of the biggest advantages of computational miRNA identification is the universal scope of the analysis; as the prediction does not require experimental material, it can potentially discover all of the miRNAs encoded by a given organism, even those that are expressed only at very low levels, in rare cells, or during brief periods of development. However, this advantage is partially annulled by the insufficient precision of the presently available algorithms: as the programs (to varying degrees) produce large numbers of false positive predictions, experimental verification is still a necessity. Northern blotting is frequently performed to investigate the expression of the computationally predicted candidates, or the predicted sequences are amplified from small RNA libraries. These procedures are not particularly compatible with high-throughput screening, and since many computational methods produce large numbers of candidates, only a small contingent of the predictions is usually subjected to experimental verification, whereas the majority remains untested. More importantly, the experimental validation methods are subject to many of the same limitations that hamper the cloning approach. Thus, even if an experimental verification is attempted and fails, it is often impossible to decide whether the failure was due to a false-positive prediction, insufficient sensitivity of the experimental detection method, or lack of expression in the tested tissue or cell line. It is thus perhaps not surprising that the expression criterion has not been satisfied for most computationally predicted miRNA candidates. While some groups have attempted to reconcile these difficulties by developing expression analysis tools that are, for example, more sensitive or allow high-throughput screening, there is also tremendous effort to increase the reliability of computational prediction methods such that experimental confirmation is becoming less important.

1.4 Computational Prediction of miRNA

A plethora of computational approaches aimed at the prediction of miRNAs have been devised, and although nearly all of them use the evaluation of features that are thought to be characteristic for miRNAs in order to identify novel candidates, they vary significantly in scope, complexity, and level of sophistication of the underlying algorithms. Some approaches strive to identify the totality of miRNAs encoded by a given organism, whereas others aim to identify only miRNAs that represent closely related ortho- or paralogous of those that are already known. Some programs investigate some of the largest genomes, those of mammals, whereas others consider only some of the smallest, those of viruses. Despite these differences, most of the approaches function according to a common scheme that might be abstracted as follows. First, a pool of input

<<小分子RNA介导的基因表达调控>>

sequences (usually representing the complete genome of a given organism) is filtered in order to limit the number of candidates that have to be evaluated by downstream algorithms. I will refer to this process as upstream filtering in the following. The filtered pool is then subjected to a structure prediction. The obtained structures are then compared to those of known pre-miRNAs, and a score calculation is performed, depending on the degree of similarity. Finally, experimental validation is attempted, usually for a selection of the highest-scoring candidates. There are considerable differences in the degree to which structural features are investigated during the scoring step; sometimes the filter might simply ensure that the candidate forms a hairpin structure, whereas in other cases it might investigate the candidate's structure down to the minutest detail. The level of sophistication, in large part, will depend on the design of the upstream filter and the efficiency with which this filter preselects a set of candidates enriched for genuine miRNAs. For example, phylogenetic conservation is the most widely used upstream filter (and at least presently, it also appears to be the most efficient). Indeed, if the sequence of a known mature miRNA is perfectly conserved in a closely (or even distantly) related species, a relatively simple structural analysis that shows that the ability of the surrounding sequences to adopt a fold-back structure is conserved as well might suffice. In contrast, an *ab initio* prediction method in which the upstream filter is minimal will require a much more detailed structural analysis during the downstream scoring step. Thus, a highly efficient upstream filter requires a less elaborate downstream structure evaluation, and vice versa. The cloning method might be considered a special case of this scheme in which the upstream filtering is based on an experimental procedure; since this method produces only little background, the subsequent structural investigation can be minimal. All of the available computational approaches are subject to the production of false-positive (i.e., candidates that pass the filters but do not represent genuine miRNAs) and false-negative predictions (i.e., bona fide miRNAs that are rejected during the upstream filtering or the downstream scoring step). The ratio with which true-positive versus false-positive predictions are made will determine the algorithm's accuracy, while the ratio of true-positive versus false-negative predictions will determine its sensitivity. Such rates are frequently estimated in order to judge an algorithm's performance. Estimating the rate of false-negative predictions is a relatively straightforward process. Often, only a limited number of the contingent of known miRNAs is used to establish the parameters of the filtering and scoring algorithms. The remaining miRNAs are then subjected to the prediction procedure, and the number of rejected versus retained miRNAs is determined. Alternatively, the full complement of miRNAs is repeatedly passed through the filters, and the method parameters are adjusted until an acceptable ratio between rejected and accepted miRNAs is achieved (what exactly an acceptable ratio is will greatly vary with the overall design and scope of the method). The estimation of false-positive prediction rates is a more complicated matter: in order to measure such numbers with high reliability, one ideally would have a set of sequences that assuredly does not contain any miRNAs at all, or a set in which all of the genuine miRNAs are known beforehand. In theory, such a set can be created artificially from randomly generated sequences, or by shuffling naturally occurring ones, but since biological sequences are nonrandom, such a reference set would be hardly representative of the experimental sequence set. Alternatively, one might select genetic elements that have known functions and are thus unlikely to additionally represent miRNAs, but this would reduce the complexity of the reference set so drastically that the gained information would be close to meaningless. In reality, the rate of false-positive predictions is often estimated on an experimental basis. For this purpose, a representative subset of the predictions (or all of

<<小分子RNA介导的基因表达调控>>

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:<http://www.tushu007.com>