

利用轨迹自动识别方法研究单细胞内表皮生长因子受体的运输*

李辉 段兆文 竇硕星 王鹏业†

(中国科学院物理研究所, 软物质物理重点实验室, 北京凝聚态物理国家实验室, 北京 100190)

(2011年3月18日收到; 2011年5月23日收到修改稿)

利用基于方向持续度及均方位移的轨迹自动识别方法, 研究了单细胞内表皮生长因子受体 (EGFR) 运输轨迹的模式特征, 将 EGFR 运动逐段划分为定向运动、超扩散运动、布朗运动和次扩散运动四种模式, 计算了不同运动模式的动力学参数并讨论了其所揭示的受体运动的生物信息.

关键词: 轨迹识别, 胞内运输, 表皮生长因子受体, 动力学特征

PACS: 87.16.ad, 87.16.Wd

1 引言

借助单分子荧光成像技术的发展, 生物体系在纳米尺度上的动态信息得到越来越广泛的揭示. 其中, 通过分析单个分子或者基团的运动轨迹来研究其动力学特征是最重要的应用之一^[1]. 例如, 通过荧光标记相应的生物分子, 对神经细胞中轴突运输以及真核细胞中胞吞运输的过程进行分析观测, 获得运动速度、扩散系数等动力学特征值^[2,3], 从而研究细胞信号传导^[4]、内吞分选^[5]、马达蛋白工作模式^[6]等多方面的生物机理. 此外, 这种单分子研究方法在病毒研究上的应用也十分有效, 通过对病毒蛋白的选择性荧光标记, 成功阐明了其侵入、运输以及释放的过程和机理^[7].

表皮生长因子受体 (EGFR) 是细胞膜上的受体蛋白, 在表皮生长因子 (EGF) 刺激下, EGFR 会发生内吞并被包裹在囊泡内转运^[8]. 其胞内运输过程中不仅有定向运动, 也有非定向的扩散运动. 定向运动源于马达蛋白如动力蛋白、驱动蛋白等水解 ATP 获得能量, 牵引着“货物”沿微管的伸展方向持续行走^[9,10]; 扩散源自分子无规则的热运动 (布朗运动); 此外, 若在马达蛋白的协助下形成方向性扩散, 则属于超扩散运动^[11]; 若扩散的空间被局限, 则属于次扩散运动^[12]. 我们所观测到的轨

迹一般包含有多种模式, 仅对整条轨迹进行分析会掩盖其中短暂运动的动态特征, 因此需要预先对整条轨迹中具有显著性特征的集群所构成的运动模式予以识别, 然后分别进行统计分析. 近年来文献中报道了多种对轨迹进行分析的方法, 比如, 均方位移 (MSD) 法^[1,13]、自组织映射法^[14]和机器学习法^[15]等. Arcizet 等^[13]将微球在细胞内的运动分为定向和扩散两类运动, 并主要对定向运动进行了分析. 但是, 扩散也有多种运动模式, 其不仅是胞内运输的一个重要途径, 而且与很多细胞功能密切相关^[2,16], 因此对细胞内的扩散运动按不同模式划分后再分别研究是十分必要的.

本文利用基于方向持续度和均方位移的轨迹自动分析方法研究了 EGFR 在单个肺癌细胞 A549 内运输轨迹的模式特征. 我们以量子点 (QD) 标记 EGF, 在荧光显微镜下实时观察 EGFR 在活细胞内运输的动态过程. 而后将所观察的 EGFR 运输轨迹分为以下四种模式: 在马达蛋白帮助下的定向运动、超扩散运动、被动的布朗运动以及次扩散运动. 通过进一步分析单细胞内的不同运动模式, 我们得到这些运动模式的分布, 以及定向运动速度、持续时间、扩散系数等动力学参数, 并讨论了其所揭示的受体运动以及细胞内微环境的生物信息. 此外, 加入微管降解药物诺考达唑 (nocodazole) 后, 定

* 国家自然科学基金 (批准号: 10834014) 和国家重点基础研究发展计划 (批准号: 2009CB930704) 资助的课题.

† E-mail: pywang@aphy.iphy.ac.cn

向运动以及超扩散运动的比例发生了明显的降低, 不仅说明了前两种运动模式都与微管相关, 而且进一步验证了我们的轨迹识别结果.

2 实验系统与分析方法

2.1 EGFR 胞内运输成像实验装置

实验所观测的是 EGFR 在 A549 肺癌细胞内的运输过程. 采用 QD (发射光波长为 655 nm) 及 EGF 来标记细胞膜上的 EGFR^[17]. 整个标记过程在冰浴上操作, 以防止在标记时发生 EGFR 的内吞. 首先对细胞中加入浓度为 10 nM 的生物素 - 表皮生长因子复合物 (biotin-EGF), 静置 10 min, 用于标记细胞表面的 EGFR. 然后用磷酸盐缓冲液 (PBS) 冲洗 3 遍, 加入浓度为 1 nM 的链霉亲和素 - 量子点复合物 (streptavidin-QD) 作用 1 min, 利用生物素与链霉亲和素之间的特异性结合的特点, 标记上一步骤中已经链接到细胞表面的 biotin-EGF. 最后用 PBS 洗 3 遍, 换入成像用细胞培养液 DMEM (不含酚红). 在药物处理实验中, 实验前加入 60 μ M 浓度的诺考达唑, 培育 30 min, 以使细胞内微管完全解聚, 而且在实验过程中该药物一直存在.

在 IX70 倒置荧光显微镜上观察 EGFR 的胞内运输过程. 为了维持细胞在正常生理状态, IX70 上安装了细胞培养观察载物台, 可以保持 37 $^{\circ}$ C 恒温及 5% 的二氧化碳浓度. 物镜放大倍率为 60, 数值孔径 NA 为 1.45; 激发光波长为 532 nm, 以最大限度减少细胞内自发荧光; 采用电子倍增 CCD (EM-CCD, DU897, Andor) 作为探测器. 在 37 $^{\circ}$ C 下, 细胞膜上的 EGFR 会被内吞到细胞内部, 并进入胞内运输通道. 我们所采集的 EGFR 运动图像位于内吞开始后 20—60 min 之间, 图像采集频率为 10 Hz. 实验后提取录像中的 EGFR 运动轨迹, 使用了软件 ImageJ 中的插件 Particles Detection and Tracking^[18].

2.2 定义相关名词

每个轨迹都是由 n 个按照图像采集顺序的目标物所在位置坐标 (x_i, y_i) 来构成, 这里将轨迹定义为 $T = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$. 由这些位置坐标 (x_i, y_i) 可以计算出每一帧的时间里目标物移动的距离, 将其定义为步长 $s_i = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}$ $i \in [1, n-1]$. 每一步所用的时间即图像的间隔定义为 τ , 在本文

里 $\tau = 100$ ms. 因为运动模式反映了一段连续轨迹的特征, 所以在模式分析时要选定所计算的轨迹点的数量, 这里定义其为窗口大小 w , 分析不同的运动模式需要选取不同的数值.

2.3 轨迹识别算法

所有计算均在 Matlab 编程环境下实现. 我们将胞内运输分为定向运动、超扩散运动、布朗运动、次扩散运动四种模式. 首先识别轨迹中的定向运动, 特点是运动方向基本保持不变. 对于轨迹上的每一个点 (x_i, y_i) , 计算它的方向持续度 ϕ_i :

$$\phi_i = \frac{1}{w} \sum_{j=i-w/2}^{i+w/2-1} \cos \theta_j, \quad (1)$$

其中 θ_i 为相邻两步 s_i 和 s_{i+1} 角度变化, 在直线运动中角度变化值为 0, 因此 $\phi_i = 1$ 表示理想的定向运动. 但是考虑到微管带有自然弯折, 以及实验图像的采集速度和误差, 我们选定 $\phi_i \geq \sigma_\phi$ ($\sigma_\phi = 0.6$) 作为判断点 (x_i, y_i) 所在区域为定向运动的标准, 并且窗口大小为 $w = 10$. 同时, 限定筛选定向运动的最短长度是 10 帧, 那么就可以将所有轨迹中满足定向运动条件的连续部分进行标记, 以提取分析.

然后识别其他运动模式, 计算轨迹上每个点 (x_i, y_i) 的均方位移 MSD:

$$\text{MSD}(k) = \frac{1}{4w - k} \sum_{i=-2w}^{2w} [(x_{i+k} - x_i)^2 + (y_{i+k} - y_i)^2], \quad (2)$$

这里 $k = 1, 2, \dots, 2w$. 鉴于扩散等非定向运动的特征需要更多的轨迹点来呈现, 为了更加准确地挑选非定向运动, 在计算 MSD 时的窗口大小为定向运动中的四倍即 $4w$. 然后用下面的式子对 MSD 进行拟合:

$$\text{MSD} = 4D(k)^\alpha, \quad (3)$$

其中 D 是扩散系数, 表示目标物的扩散能力, 与扩散颗粒的半径和介质的黏滞系数正相关. 指数 α 反映了此点所在区域的运动模式^[1], 我们定义其为模式系数. 当 $\alpha = 2$ 时, 代表理想的定向运动; $\alpha = 1$, 代表了布朗运动; 而当 $1 < \alpha < 2$ 时, 运动特点介于定向运动与布朗运动之间, 我们把这种有方向的扩散过程定义为超扩散运动; 当 $0 < \alpha < 1$ 时, 则表示运动是在一种受限区域内的扩散过程, 定义为次扩散运动.

3 结果与讨论

我们通过以上轨迹识别方法处理了实验中所观察的 EGFR 运动轨迹, 并将轨迹划分为定向运动、超扩散运动、布朗运动、次扩散运动这四种模式. 图 1(a) 是一条典型的运动轨迹, 利用算法识别后的结果在图中做了标记以便区分, 它含有多段定向运动 (粗实线)、超扩散运动 (虚线) 和次扩散运动 (细线), 而布朗运动很少 (未标记). 通过观察图 1 中的轨迹形态, 可以看出我们所识别出的轨迹模式已经达到了人工挑选的效果, 表明我们的算法是准确可靠的. 图 1(b) 详细说明了图 1(a) 中轨迹的前 60 s (阴影部分) 的识别过程. 通过计算每个轨

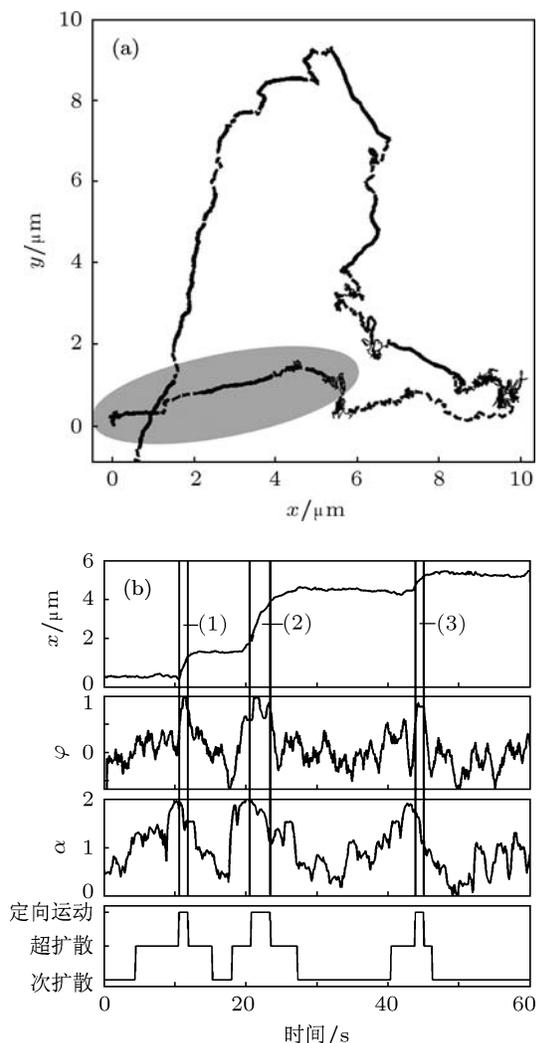


图 1 胞内运输轨迹的识别 (a) 典型轨迹识别结果, 粗实线代表定向运动, 虚线代表超扩散运动, 细线代表次扩散运动, 布朗运动数量少不易观察, 所以未标记; (b) 详细说明 (a) 图轨迹前 60 s (阴影内) 的识别过程, 图片从上至下为: x 方向上的位置、方向持续度 ϕ 、模式系数 α , 以及划分的运动模式相对于时间的曲线图, 图中标记了三段定向运动

迹点的方向持续度 ϕ 和扩散系数 α , 并考察其是否满足所定义的轨迹划分条件, 最终这段轨迹被划分成三种运动模式. 图中有三段被识别出来的定向运动, 与 x 轴方向的移动曲线图可以准确地符合, 说明此算法可以准确地挑选出此运动模式.

我们认为所识别的定向及超扩散运动不仅需要马达蛋白的介入, 同时微管作为运输的“道路”也起到至关重要的作用. 为了检验所挑选的上述两种运动模式与微管的相关性, 我们做了相关药物对照实验: 在实验前对细胞加入一种常用的微管解聚药物——诺考达唑, 作用 30 min 后使细胞内的微管发生解聚, 然后观察 EGFR 的胞内运输. 通过分析正常细胞和药物处理细胞中的大量运动轨迹可以发现: 在诺考达唑处理细胞后, EGFR 的运动性发生了明显的降低, 表现为 α 值集中分布于 0—1 之间 (见图 2), 而且平均值由正常细胞中的 0.798 下降为药物处理后的 0.496. 可见, 解聚微管抑制了模式系数 $\alpha > 1$ 的运动, 即定向和超扩散运动两种模式, 这不仅表明它们都是微管上的主动运输过程, 更进一步验证了算法的准确性和可靠性.

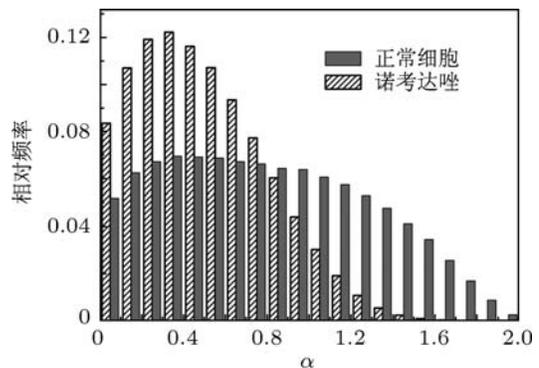


图 2 正常细胞和诺考达唑处理的细胞中模式系数 α 的分布

我们分别统计了在正常和诺考达唑处理的细胞中四种运动模式的比例 (表 1). 可以看出, 定向运动在正常细胞中所占的比例很低 (1.12%), 然而在诺考达唑处理的细胞中其几乎被完全抑制, 其比例仅为 0.005%; 超扩散运动的比例在微管解聚后由 29.3% 降为 3.59%, 变化也十分明显. 而布朗运动在两种细胞中的比例都较低, 而且数值接近. 其比例低的原因可能是因为细胞质内有多种细胞骨架及大量的细胞器, 其内部环境十分拥挤, 所以含有 EGFR 的囊泡没有足够的空间做自由布朗运动. 此外, 囊泡的布朗运动与微管无关, 所以它所占的比例不受药物处理的影响. 次扩散运动在正常细胞中的比例高达 61.9%, 而在微管解聚的细胞中, 其上

升至 91.6%, 说明在复杂粘稠的细胞质中, 囊泡大部分运动局限在一个区间里, 在马达蛋白和微管的帮助下才能够实现细胞内的有效转运. 同时表明只有

预先对轨迹进行识别分类, 才能准确的获得所关注的运动模式的动态信息, 否则它的特点会被更多的次扩散运动所掩盖.

表 1 正常细胞和诺考达唑处理后的细胞中四种运动模式的比例统计

		定向运动	超扩散运动	布朗运动	次扩散运动
正常细胞 (总数 1177192)	比例 (%)	1.12	29.3	7.61	61.9
	样本数	13869	345395	89546	728382
诺考达唑处理 (总数 1484761)	比例 (%)	0.005	3.59	4.77	91.6
	样本数	20	53336	70865	1360540

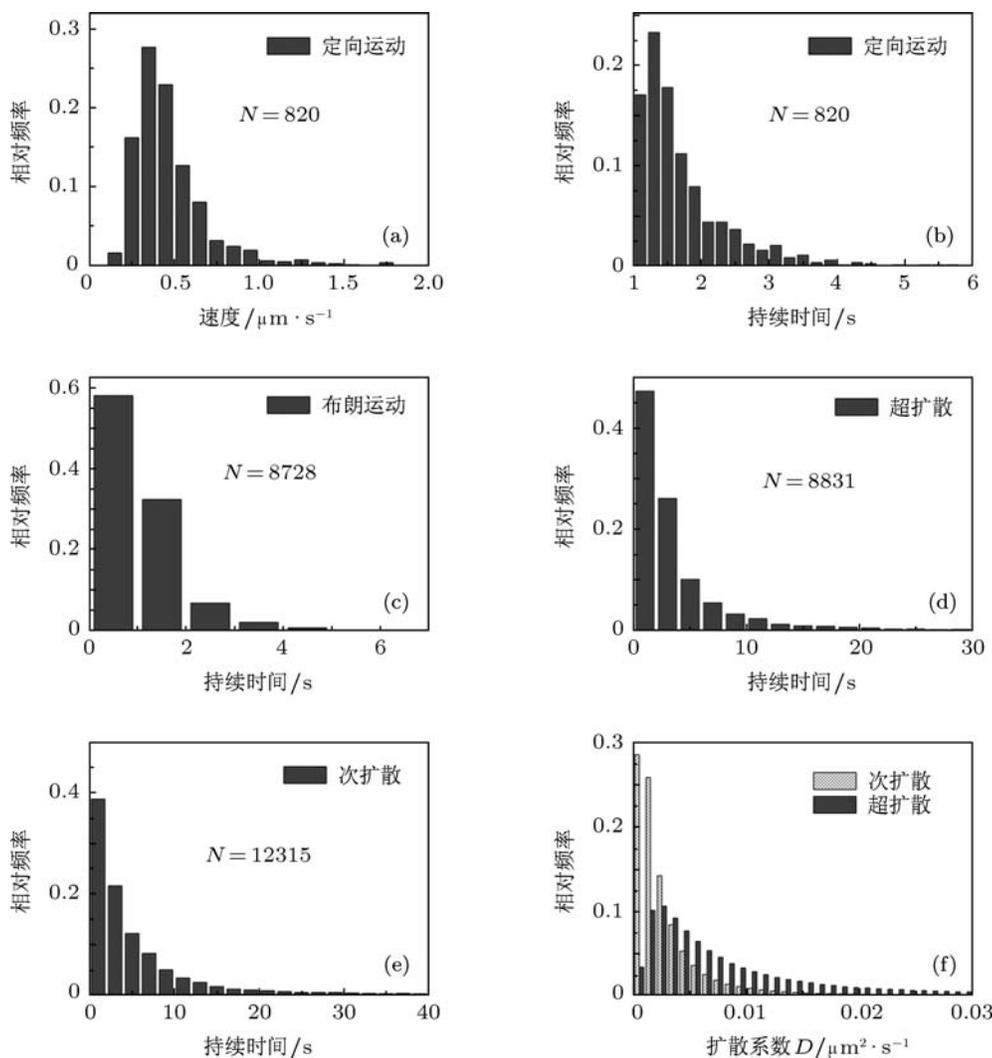


图 3 统计四种运动模式的动力学特征 (a) 定向运动的速度分布; (b) 定向运动的持续时间分布; (c) 布朗运动的持续时间分布; (d) 超扩散运动的持续时间分布; (e) 次扩散运动的持续时间分布; (f) 超扩散和次扩散运动的扩散系数分布

接下来, 我们分别统计了正常细胞中上述四种运动模式的动力学特征. 在图 3(a) 和 (b) 中可以看出, 定向运动 (轨迹数 820) 的速度分布比较宽 (0.2—1.0 $\mu\text{m/s}$), 可能与马达蛋白的协同作用有关, 其平均速度 (0.473 $\mu\text{m/s}$) 与以往文献里报道的结果一致 [6,19]; 另一方面, 定向运动的持续时间很

短, 平均值为 1.69 s, 表明马达蛋白在微管上持续行走 1—2 s 后就会脱落下来, 这与马达蛋白的工作机理有关 [10]. 布朗运动 (轨迹数 8728) 的持续时间更短 (见图 3(c)), 其平均值仅为 1.03 s, 考虑到细胞中非常拥挤的内部环境, 这个时间尺度是合理的. 然而超扩散运动 (见图 3(d), 轨迹数 8831)

和次扩散运动 (见图 3(e), 轨迹数 12315) 的持续时间都比较长, 其平均值分别为 3.91 和 5.91 s. 对于这两种运动模式, 比较了它们的扩散系数 D , 由图 3(f) 可见超扩散运动的扩散系数分布广, 其平均值为 $0.017 \mu\text{m}^2/\text{s}$, 而次扩散运动的扩散系数主要分布在小于 $0.01 \mu\text{m}^2/\text{s}$ 的区域, 其平均值仅为 $0.004 \mu\text{m}^2/\text{s}$. 这个结果验证了以往报道的甘氨酸受体在突触中的快速扩散系数 ($0.053 \mu\text{m}^2/\text{s}$) 和慢速扩散系数 ($0.002 \mu\text{m}^2/\text{s}$)^[2].

4 结论

利用基于方向持续度和均方位移的轨迹自动

识别方法研究了 EGFR 在单细胞内的运输过程, 将其划分为定向运动、超扩散运动、布朗运动和次扩散运动四种模式. 通过分析不同的运动模式, 并比较了诺考达唑解聚细胞内微管后的实验结果, 不仅证明了该识别方法的可靠性, 而且获得了更准确的受体运动的动力学参数及其所揭示的生物信息. 说明本轨迹识别方法能够为进一步了解细胞体内生物分子在纳米尺度上的运动传输提供帮助. 当然, 特殊的运输体系可能会影响轨迹的识别结果, 如神经细胞中长距离的轴突运输、病毒对细胞的多阶段侵染过程等都表现出不同的运动特征. 因此, 我们下一步要对这些问题进行更深入的研究.

-
- [1] Saxton M J, Jacobson K 1997 *Annu. Rev. Biophys. Biomol. Struct.* **26** 373
- [2] Dahan M 2003 *Science* **302** 442
- [3] Cui B X, Wu C B, Chen L, Ramirez A, Bearer E L, Li W P, Mobley W C, Chu S 2007 *Proc. Natl. Acad. Sci. U. S. A.* **104** 13666
- [4] Gonzalez-Gaitan M 2003 *Nat. Rev. Mol. Cell Biol.* **4** 213
- [5] Lakadamyali M, Rust M J, Zhuang X 2006 *Cell* **124** 997
- [6] Watanabe T M, Higuchi H 2007 *Biophys. J.* **92** 4109
- [7] Brandenburg B, Zhuang X 2007 *Nat. Rev. Microbiol.* **5** 197
- [8] Sorkin A, Goh L K 2009 *Exp. Cell Res.* **315** 683
- [9] Kural C, Kim H, Syed S, Goshima G, Gelfand V I, Selvin P R 2005 *Science* **308** 1469
- [10] Ping X, Dou S X, Wang P Y 2004 *Chin. Phys.* **13** 1569
- [11] Vale R D 2003 *Cell* **112** 467
- [12] Kusumi A, Nakada C, Ritchie K, Murase K, Suzuki K, Murakoshi H, Kasai R S, Kondo J, Fujiwara T 2005 *Annu. Rev. Biophys. Biomol. Struct.* **34** 351
- [13] Arcizet D, Meier B, Sackmann E, Radler J O, Heinrich D 2008 *Phys. Rev. Lett.* **101** 248103
- [14] Owens J, Hunter A 2000 *VS '00 Proceedings of the Third IEEE International Workshop on Visual Surveillance (VS'2000)* Dublin, Ireland, July 1, 2000 p77
- [15] Helmuth J A, Burckhardt C J, Koumoutsakos P, Greber U F, Sbalzarini I F 2007 *J. Struc. Biol.* **159** 347
- [16] Verkman A S 2002 *Trends Biochem. Sci.* **27** 27
- [17] Jaiswal J K, Mattoussi H, Mauro J M, Simon S M 2003 *Nat. Biotechnol.* **21** 47
- [18] Sbalzarini I F, Koumoutsakos P 2005 *J. Struc. Biol.* **151** 182
- [19] Courty S, Luccardini C, Bellaiche Y, Cappello G, Dahan M 2006 *Nano Lett.* **6** 1491

The study of EGFR transport in single cell using an automatic method of trajectory identification*

Li Hui Duan Zhao-Wen Dou Shuo-Xing Wang Peng-Ye[†]

(Key Laboratory of Soft Matter Physics, Beijing National Laboratory for Condensed Matter Physics, Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China)

(Received 18 March 2011; revised manuscript received 23 May 2011)

Abstract

We study the characteristics of motion modes of epidermal growth factor receptor (EGFR) trafficking in single cell, by using the automatic method of trajectory identification which is based on the directional persistence and mean square displacement analysis. Each trajectory of EGFRs is divided into four modes of motion: directional motion, super-diffusion, Brownian motion and sub-diffusion. The corresponding dynamic parameters of different modes of motion are calculated and discussed, providing an insight into intracellular trafficking of receptors.

Keywords: trajectory identification, intracellular transport, epidermal growth factor receptor, dynamic characteristics

PACS: 87.16.ad, 87.16.Wd

* Project supported by the National Natural Science Foundation of China (Grant No. 10834014) and the National Basic Research Program of China (Grant No. 2009CB930704).

[†] E-mail: pywang@aphy.iphy.ac.cn