

# 低氧反应途径促进线虫 PEP 羧激酶和糖异生

生信 2001 张子栋 翻译

2023 年 3 月 22 日

## 摘要

活跃的细胞分裂，包括一些癌症，依靠有氧糖酵解而不是氧化磷酸化来产生能量，这种现象被称为 Warburg 效应。低氧诱导因子 (HIF-1) 的结构性激活是 Warburg 效应的一个标志，HIF-1 是一种转录因子，以介导对缺氧的适应性反应而闻名。

HIF-1 被认为可以促进糖酵解和抑制氧化磷酸化。在这里，我们相反地展示了 HIF-1 可以促进糖异生。利用多组学方法，我们揭示了线虫基因组、转录和代谢环境受构成活性的 HIF-1 调控。我们使用 RNA-seq 和 ChIP-seq 在有氧条件下分析缺失 HIF-1 关键负调控因子 *egl-9* 的突变株。我们整合了这些方法来识别 200 多个直接和功能上受 HIF1 上调的基因，包括糖异生的限速介质 PEP 羧基激酶 PCK-1。这种由 HIF-1 激活的 PCK-1 促进了对氧化和低氧应激的生存。我们的工作确定了体内 HIF-1 的功能直接靶点，全面描述了 HIF-1 激活在生物体中诱导的代谢组。

# 目录

<b>1</b>	<b>引言</b>	<b>1</b>
<b>2</b>	<b>结果</b>	<b>2</b>
2.1	确定 HIF-1 的直接靶点 . . . . .	2
2.2	HIF-1 结合位点的特征 . . . . .	3
2.3	HIF-1 对新陈代谢的再编程 . . . . .	4
2.4	PCK-1 促进低氧应激时的存活 . . . . .	5
2.5	HIF-1 直接靶点是保守的 . . . . .	6
<b>3</b>	<b>讨论</b>	<b>7</b>
<b>4</b>	<b>方法</b>	<b>9</b>
4.1	转基因和转基因动物的产生 . . . . .	9
4.2	荧光显微镜与图像分析 . . . . .	10
4.3	定量 RT-PCR 检测基因表达 . . . . .	10
4.4	产卵和卵滞留的特征 . . . . .	11
4.5	ChIP-seq . . . . .	11
4.6	RNA-seq . . . . .	13
4.7	用 BETA basic 和 LOA 识别直接靶标 . . . . .	14
4.8	Motif 识别和富集 . . . . .	14
4.9	HIF-1 结合位点转录因子占位的鉴定 . . . . .	15
4.10	代谢组学 . . . . .	15
4.11	百草枯存活试验 . . . . .	15
4.12	低氧存活试验 . . . . .	15

## 1 引言

缺氧在包括缺血性中风、心肌梗死、肺动脉高压、脑瘫、新冠肺炎和癌症在内的多种人类疾病中起着核心作用。除了通过损害氧化磷酸化 (OXPHOS) 来减少能量 (ATP) 的产生外, 缺氧和随后的复氧也通过产生有毒的活性氧 (ROS) 来产生氧化应激。在涉及急性缺氧的疾病中, 能量缺乏和氧化应激都会导致组织损伤。

后生动物通过一条保守的反应途径对缺氧作出反应。在有氧 (常氧) 条件下, 脯氨酸羟化酶感知氧气, 并用它来共价修饰和负调控缺氧诱导因子 (HIF  $\alpha$ ), 这是该途径的转录效应 (图 1a)。当随后发生缺氧时, 由于缺氧, 脯氨酸羟化酶被抑制, 导致缺氧诱导因子  $\alpha$  的去抑制。

如果缺氧诱导因子  $\alpha$  变得稳定, 就会与 HIF $\beta$  二聚化, 在整个基因组中调节特定靶基因的转录。HIF 的激活可以将急性缺氧造成的损害降到最低; 然而, 由缺氧反应途径本身引发的长期适应和组织重塑可能是有害的。

缺氧诱导因子通过调节多种生理过程来抵消缺氧性损伤。HIF 上调介导葡萄糖摄取和糖酵解的酶, 以及丙酮酸脱氢酶激酶, 丙酮酸脱氢酶激酶是代谢产物进入三羧酸 (TCA) 循环的负调节因子。HIF 通过改变电子传递链 (ETC) 的亚单位组成和减少线粒体数量来抑制 OXPHOS。由此产生的代谢开关优化了细胞在无氧条件下产生 ATP 的能力。在某些情况下 (例如, 癌症、分裂干细胞、激活的 T 淋巴细胞、子宫内膜蜕膜化, 假设患者正在接受 Pro 羟化酶抑制剂治疗贫血), HIF 激活这一开关并促进糖酵解, 尽管有氧条件 (即 Warburg 效应)。HIF 还调节目标基因, 如 EPO 和血管内皮生长因子, 这在发育中重塑氧气输送。我们才刚刚开始发现 HIF 靶点的完整列表, 确定这些靶点中哪些是直接的, 哪些是间接的, 并展示哪些与活体相关。

我们对缺氧反应的大部分了解来自对培养细胞的研究。在这里, 我们使用遗传模式生物线虫来描述体内 HIF 激活诱导的适应性反应。这些线虫采用保守的低氧反应途径的单一同源基因 (Pro 羟化酶同源基因 *egl-9* 和低氧诱导因子 同源基因 *hif-1*), 在 *hif-1* 和 *egl-9* 中有活性的零突变体对低氧胁迫表现出不同的敏感性。线虫的低氧反应途径除了调节低氧下的存活外, 还调节行为、神经元中谷氨酸受体的转运、线粒体动力学、产卵回路和寿命。HIF-1 在 *egl-9(Sa307)* 突变体中具有结构性活性, 我们在这里使用基因组、转录和代谢组学的综合方法来获得完整动物中 HIF 激活的整体和功能描述。虽然 HIF 在促进糖酵解方面的作用得到了很好的认识,

但我们发现 HIF-1 也通过激活 PCK-1 PEP 羧基激酶来促进糖异生和抗氧化剂的产生。HIF-1 通过 PCK-1 发挥作用，促进对氧化和低氧应激的存活。我们的工作确定了体内 HIF-1 的直接功能靶点，全面描述了 HIF-1 激活在生物体中诱导的代谢组。

## 2 结果

### 2.1 确定 HIF-1 的直接靶点

为了确定与 HIF-1 结合的基因组中的位点，我们创建了一个 *odIs131[hif-1::gfp]* 转基因，并将其表达嵌合 HIF1::GFP 到 *hif-1(ia4)* 缺失突变体中，发现它的表达水平与内源性 HIF-1 相似 (补充图 1a-e)。

在有氧条件下，使用 *egl-9* 缺失的突变背景来激活 HIF-1 (图 1a)，我们观察到来自 *hif-1(ia4); odIs131[hif-1::gfp]* 动物的少量荧光 (图 1b)，但在 *egl-9* 突变的几乎所有组织中都有明显的核 HIF-1::GFP 荧光 (图 1c)。此外，*odIs131[hif-1::gfp]* 在多个突变表型 (补充图 1f-i) 中取代了内源性 HIF-1，为鉴定生理条件下的基因组结合位点提供了一种功能工具。

为了确定 HIF-1 直接调控的基因，我们使用抗 GFP 抗体和高通量测序对 L4 期 (生育前期) *egl-9(sa307); hif-1(ia4); odIs131[hif-1::gfp]* 动物进行了 ChIP-seq。

我们确定了与 HIF-1::GFP 结合位点 (补充数据 1) 相对应的 604 个测序读取峰，其中大多数落在附近基因的 500 bp 以内，并对人类缺氧反应元件 (HRE，补充图 2a-c) 进行了浓缩。

最接近芯片序列峰的基因并不总是受相关转录因子结合位点调控的直接靶点。因此，我们使用 RNA-SEQ 技术来分析由于 HIF-1 激活引起的差异基因表达，利用这些数据来确定 HIF-1 结合位点附近受 HIF-1 调控的基因。我们研究了 L4 期动物在四种不同的成对基因组合中的转录，其中 HIF-1 活跃和不活跃，考虑了两个实验比较，这两个比较将丰富 HIF-1 调节的差异表达基因 (DEG)： *egl-9* 突变体 (活跃的 HIF-1) 和野生型 (非活性 HIF-1) 和 *egl-9* 突变体 (活性 HIF-1) 与 *egl-9* HIF1 双重突变体 (非活性 HIF-1) (图 1d 和补充图 3a, b)。对于这两个实验比较，我们利用 *odIs131[hif-1::gfp]* 转基因来拯救 HIF-1 突变体，以提供额外的实验数据集，减少不相关的遗传背景影响。我们的成对方法确保比较的交集包含高严格的 HIF-1 调节基

因 (补充图 3c, d 和补充数据 2)。

从不同分析的交集 (重叠) 确定 DEG 的一个缺点是, 对每个分析的任意阈值的依赖可能导致低估重叠的大小。因此, 我们使用 Luperchio 重叠分析 (LOA)<sup>40</sup> 来识别 DEG, 在没有 *odIs131* 转基因的成对比较中使用潜在 DEG 的证据来通知包含转基因的成对比较中那些潜在 DEG 的状态 (图 1d)。我们推测, 与野生型 (LOA 组合 1) 相比, *egl-9* 突变株与野生型 (LOA 组合 1) 相比, 在 *egl-9* 突变体中富含最高可信度的 HIF-1 调控的 DEG; 与 *egl-9 hif1* 双突变体 (LOA 组合 2) 相比, *egl-9* 突变体中的 DEGS 将富含。我们用 beta 软件<sup>41</sup> 分析了得到的 DEG 列表, 该软件通过整合 DEGS 和 ChIP-seq 数据推断出直接靶基因, 以确定 360 个附近 HIF-1 ChIP-seq 结合位点的 216 个差异表达的直接靶基因 (图 1d 和补充数据 3)。当 HIF-1 激活时, 每个识别的直接靶点都被独占地上调, 与转录激活剂一致 (图 1e)。直接靶被浓缩为 HRE 序列 (图 1f)。我们确定了先前被证明受 HIF-1 调控的基因, 进一步验证了我们的方法 (图 1g)。超过一半 (59%) 的 HIF-1 靶标与单个结合位点相关, 表达的幅度与其结合位点的数量松散相关 (图 1g)。HIF-1 的两个关键负调控因子 (*rhy-1* 和 *egl-9*) 包含一些最高数量的 HIF-1 结合位点, 并被 HIF-1 显著上调, 表明作为反应的一部分有一个强大的负反馈环 (补充图 3e)。

## 2.2 HIF-1 结合位点的特征

为了确定 HIF-1 对这些靶基因的结合位点是否具有功能, 我们构建了荧光 Venus 转录报告基因 (图 2a 和补充图 4a), 用于一个新的靶基因 *pck-1* (PEP 羧化酶) 和一个已知的靶基因 *rhy-1*。在 *egl-9* 突变体中, HIF-1 是活跃的, 从转录报告基因中观察到 Venus 荧光水平升高, 与野生型和 *egl-9; hif-1* 双突变体相比, HIF-1 是不活跃的 (图 2b-g 和补充图 4b)。在转录报告基因中, 如果完整的编码 HIF-1 ChIP-seq 峰值的序列被移除 ( $\Delta$  峰), 或者峰值中的 6 bp 核心 HRE 被删除 ( $\Delta$ HRE), 或者只有最小启动子序列存在, 则报告表达降低, 与野生型启动子相比 (图 2h 和补充图 4b), 表明 HREs 是 HIF-1 在体内调节靶基因表达所必需的。

ModENCODE/ModRED 联盟以前发现了被多个 (>15) 转录因子占据的高占有率靶点 (HOT); 这种现象可能是由于 HotSite 是基因组的“粘性”区域, 导致芯片序列数据中的错误信号。我们调查了 ModENCODE/MODEM 数据库中所有转录因子的已知结合位点与 HIF-1 位点的重叠。对所有 604

个 HIF-1 结合位点的研究发现, 大多数位于已知 HOT 位置或附近 (补充图 4c), 在最近邻基因中观察到的差异调节很少 (补充图 5a)。相比之下, 用 LOA 和 BETA 鉴定的功能 HIF-1 位点中只有大约一半位于 HOT 位点 (补充图 4c), 而与其他 HOT 位点相关的那些基因显示了 HIF-1 依赖的基因表达变化, 表明它们是功能的 (补充图 5b-e)。

小提琴图 (图 2I) 表明, 使用 LOA、BETA 和严格的 RNA-Seq 数据来识别受 HIF-1 位点调控的靶标导致了靶标调用准确性的提高, 即使只使用 BETA, 也能从 HOT 位点转移到低占有率位点 ( $<5$  其他因素限制)。像预期的 HIF-1 二聚化伴侣 3 一样, AHA-1 结合位点在 HIF-1 位点附近丰富 (图 2j)。促进抗氧化反应的线虫 Nrf2 同源基因 SKN-1 也丰富, 这表明缺氧和抗氧化反应之间存在共同的靶点。相比之下, DAF-16/FOXO 的锌指转录拮抗剂 *pqm-1* 结合的位点在 HIF-1 位点附近表达不足, 这与 *pqm-1* 突变体 45 的缺氧存活率增加一致。最后, LOA 和 BETA 的使用使得能够识别作用于受调控靶基因的 TSS 远端的调控位点 (补充图 2a, b)。

### 2.3 HIF-1 对新陈代谢的再编程

对于 HIF-1 上调的 216 个直接靶点, 糖酵解、糖异生、氨基酸代谢、硫氧化、脂肪酸  $\beta$  氧化和氧化还原代谢等 GO 术语显示丰富 (表 1), 表明 HIF-1 通过直接与代谢途径关键酶的启动子结合来重新编程代谢。

我们获得了代谢组谱, 以确定当 HIF-1 在有氧条件下活跃时, 与它不活跃时相比, 转录上调的途径中是否有更多来自这些途径的代谢物。我们观察到 175 种代谢物水平的差异 ( $P < 0.05$ ), HIF-1 激活导致各种氨基酸、碳水化合物、脂肪和核苷酸水平升高, 这表明多条代谢途径发生了变化 (图 3A 和补充数据 4)。

为了在特定生物化学途径的背景下检查代谢物水平, 我们将野生型和 *egl-9* 突变代谢组数据与催化涉及每个特定代谢物的反应的酶的相应 RNA-SEQ 数据进行了映射。正如预期的 46, 47, HIF-1 直接促进了多种关键的糖酵解酶的表达, 并增加了这一途径的代谢物种群 (图 3B)。大多数 TCA 循环酶和代谢物在 *egl-9* 突变体中没有变化 (图 3C)。然而, *egl-9* 突变体的 mdh-1(苹果酸脱氢酶) 和 LDH-1(乳酸脱氢酶) 及其相关代谢物苹果酸、丙酮酸和乳酸水平升高 (图 3D)。苹果酸和苹果酸脱氢酶的升高表明乙醛循环的活性, 乙醛循环是一些生物体用来将柠檬酸转化为草酰乙酸酯以满足糖异生途径的三羧酸循环的变体。

我们还观察到糖异生增加的指标 (图 3B)。正如前面讨论的, HIF-1 直接和显著地促进 PCK-1(PEP 羧酸激酶) 的表达, PCK-1 催化糖异生的关键限速步骤, 在该步骤中草酰乙酸酯 (OA) 被转化为磷酸烯醇式丙酮酸 (PEP)。除了再生葡萄糖外, 糖异生还为磷酸戊糖途径 (PPP) 提供底物, 该途径产生还原等量的 NADPH、5-碳糖 (例如, 用于合成核酸的核糖-5-磷酸) 和红-4-磷酸 (用于合成芳香族氨基酸)。我们观察到 *egl-9* 突变体中 PPP 代谢物的增加相对于野生型 (图 3e), 尽管 PPP 中的关键酶都不是 HIF-1 的直接靶标, 这表明 PPP 通量的增加是 HIF-1 激活糖异生的间接影响。

由 PPP 产生的 NADPH 作为脂肪酸合成的还原当量, 合成的和饮食中的脂肪酸都附着在 3-磷酸甘油 (G3P) 上, 转化为储存脂。与这一代谢途径一致, 我们观察到当 HIF-1 活跃时, G3P 减少和某些脂肪酸水平增加 (图 3f)。磷脂也是由 G3P 合成的, 我们发现活性 HIF-1 增加了某些磷脂物种的水平 (图 3f)。

NADPH 还补充还原型谷胱甘肽, 这是一种主要的抗氧化剂, 用于对抗氧化应激。我们观察到 *egl-9* 突变体中谷胱甘肽的水平高于野生型 (图 3g)。事实上, 许多由 HIF-1 最显著上调的直接靶基因位于产生谷胱甘肽的途径中 (图 3g), 而对抗氧化应激的过氧化氢酶和超氧化物歧化酶都间接上调 (图 3h)。氧化应激反应途径也有助于对抗线虫中的细菌感染, 因为铜绿假单胞菌等细菌病原体会产生 HCN 和  $H_2S$  等毒素, 这些毒素可以被这些途径中和。事实上, 我们观察到 HIF-1 直接促进整个  $H_2S/HCN$  解毒途径的表达 (图 3i), 包括 *cysl-2* (氰丙氨酸合成酶) 和 *sqrd-1* (硫代酮氧化还原酶), 它将这些毒素转化为多硫化物和硫酸盐。HIF-1 对  $H_2S/HCN$  解毒作用的上调与 *egl-9* 和 HIF-1 突变体分别对假单胞菌感染和硫化物/氰化物毒性敏感一致。

## 2.4 PCK-1 促进低氧应激时的存活

传统上, HIF-1 被认为是通过促进无氧 ATP 合成来对抗缺氧。我们的多组学分析也强调了它在 *egl-9* 突变体中促进糖异生和氧化应激反应。为了测试直接 *hif-1* 靶标是否也被缺氧上调, 我们使用 qRT-PCR 来测量野生型或 HIF-1 突变体中的 mRNA 水平 (图 4a 和补充图 6)。所有测试的靶基因都以 HIF-1 依赖的方式被缺氧上调, 包括糖异生的限速酶 *pck-1*。

鉴于缺氧和随后的复氧都导致活性氧 (ROS) 和氧化应激的产生, 我们推断 HIF-1 上调 PCK-1 也可能通过糖异生和 PPP 动员抗氧化反应来对抗



缺氧。为了排除用于培养线虫的活大肠杆菌的代谢贡献，我们用甲醛固定细菌培养物，使其代谢惰化。将这些动物置于固定的大肠杆菌上，直到它们达到 L4 阶段，然后将这些动物暴露于没有食物的低氧环境中 48 小时，测量在常氧环境下恢复 24 小时后的存活率。*pck-1* 的突变体，即 *hif-1* 突变体，在常氧条件下是存活的 (补充图 7a)。

然而，与野生型相比，*hif-1* 和 *pck-1* 单一突变体对缺氧都显著敏感 (图 4b)，并且这种易感性通过补充 PEP(PCK-1 催化的反应的产物，当 HIF-1 有活性时，该产物升高，如图 3b 所示) 而不是上游代谢物丙酮酸而得以挽救。我们观察到胚胎孵化的缺氧存活试验的可比结果 (图 4c 和补充图 7b)。抗氧化剂乙醇酸盐和 N-乙酰半胱氨酸 (NAC) 在喂给秀丽隐杆线虫时可以降低氧化应激，并且我们发现补充这些抗氧化剂拯救了缺氧期间存活的 *hif-1* 和 *pck-1* 突变体 (图 4d, e 和补充图 7c)。用生长在活大肠杆菌上的线虫观察到类似的结果 (补充图 7d, e)。此外，我们发现 *pck-1* 突变体对氧化应激敏感，因为它们在暴露于超氧化物发生剂百草枯时表现出较差的存活率 (补充图 7f)。*pck-1* 的突变没有抑制 *egl-9* 突变体的卵滞留或延长寿命表型 (补充图 8)，表明 PCK-1 的上调没有介导 HIF-1 的所有功能。总之，我们的结果表明 HIF-1 直接促进糖异生和氧化应激抗性有助于动物在缺氧状态下存活。

## 2.5 HIF-1 直接靶点是保守的

我们为线虫分析中确定的直接靶标确定了假定的人类序列同源序列 (补充数据 5)。我们重新分析了从经历 HIF1A 激活的人类细胞中公布的 RNA-seq 数据，并确定了几组显示出相关的共表达模式的同源基因 (补充图 9)。虽然只有 0.3% 的人类基因受到 HIF1A 的一致上调，但 6.8% 的针对线虫 HIF-1 直接靶点的人类同源基因显示出一致的上调，增加了 23 倍 ( $P\text{-value} < 1.0 \times 10^{-16}$ , 假设检验)。在几乎所有的研究中，HIF1A 上调了一组同源基因的表达，GO 富集与观察到的线虫对应基因相似 (图 5)。其他星系团在实验的子集中表现出上调。事实上，线虫 *pck-1* 的同源基因 PCK2 属于这样一个簇，其中还包括参与半胱氨酸和谷胱甘肽代谢的基因，这表明这种 PEP 羧基激酶是一个关键的 HIF1A 调控靶点，取决于组织或环境。

### 3 讨论

应激反应途径协调多个靶基因的调节表达，这些基因是重新平衡生理动态平衡所需的。为了满足机体在低氧环境中的能量需求，HIF 通过促进糖酵解来促进 ATP 的合成。然而，体内 HIF 靶点的性质和作用尚不清楚。我们的研究结合了 ChIP-seq 和 RNA-seq 来确定线虫中直接的 HIF-1 靶点的全序列，并结合代谢组学来了解它们的表达导致的生理变化。HIF-1 对新陈代谢进行重新编程，而不是像以前认为的那样简单地增强糖酵解。我们在这里表明，它直接促进脂肪代谢、乙醛酸循环、糖异生、PPP、 $H_2S/HCN$  解毒和谷胱甘肽合成。这项研究结合了基因组学、转录学和代谢学的方法来表征体内 HIF 激活的作用。

识别像 HIF-1 这样的转录因子的真正功能靶点的一种简单方法是将结合位点的鉴定与转录组学分析相结合，通常是通过比较给定转录因子活性和非活性的条件来进行的。将来自 ChIP-seq 分析的最近邻近基因列表与来自 RNA-seq 分析的差异表达基因 (DEG) 列表进行比较，这些列表的交集被认为是功能靶点。然而，仅仅基于芯片序列结合数据和 TSS 邻近来指定一个基因作为转录因子的靶标是有局限性的。这种“最近邻”方法并不评估针对每个假定靶点的转录因子活性的表达变化。这个问题在拥有紧凑基因组的生物体中变得复杂，比如线虫或果蝇，通常在一个给定的芯片序列峰附近发现多个基因。因此，这种方法确定的任何目标可能不是真正的功能目标。

一种更有效的方法是使用像 BETA 算法这样的软件，该算法使用来自 RNA-seq 研究的 DEG 模式，比较 TF 活跃和不活跃时的 mRNA 水平，以确定给定 TF(从 ChIP-seq 确定) 结合部位附近的哪个基因是靶标。除了邻近之外，BETA 还优先考虑功能表达的差异，从而增强其识别真正功能靶点的能力。这种方法虽然严格，但可能会导致对重叠大小的低估，因为必须为每个单独的分析分配任意的重要性阈值或等级。为了克服这一缺陷，我们将 BETA 与重叠分析方法 (Luperchio 重叠分析或 LOA40) 相结合，允许多次比较相互提供关于差异基因表达的证据，以克服这一限制。这种分析的理由是，在给定条件下的一组 DEG(例如，野生型与 *egl-9* 突变体) 是关于相同基因在另一类似条件下的状态的信息 (例如，野生型与 *egl-9* 突变体，其也包含 HIF-1 突变和挽救 HIF-1::GFP 转基因)。通过使用这种方法，与使用更严格的交叉方法相比，我们能够扩大功能目标的数量。我们的组合方法允许识别受调控基因 TSS 更远的 HIF-1 结合位点 (补充图 2a, b)。

鉴定功能性转铁蛋白结合位点的另一个复杂因素是高占有率靶 (HOT)

位点的存在, 这是在分散在基因组 43 上的多个独立 ChIPseq 研究中发现的过度代表的结合峰。大多数 HOT 位点要么是芯片序列伪影, 要么是结合位点, 没有明显的功能作用, 至少在调节邻近基因的转录方面是这样。在这里, 我们的组合测试版和 LOA 方法再次克服了这一限制。我们表明, 由 BETA 和 LoA 确定的目标基因更具特异性, 因为与单独的 BETA(交集) 和最近邻方法 (图 2I) 相比, 归因于基因组内 HOT 位置的目标基因的数量减少了。此外,  $\beta$  和 LOA 能够区分功能 HOT(即, 具有邻近基因的位点, 显示 HIF-1 依赖的差异调控) 和非功能 HOT(补充图 5)。最近邻方法鉴定的非功能 HOT 附近的基因没有显示任何 GO 项浓缩 (Panther 分析, FDR<0.05)。相比之下, 由 LOA 鉴定的功能 HOT 附近的基因表现出与非 HOT 附近基因类似的 GO 富集, 包括糖酵解、糖异生和氨基酸代谢, 进一步证明了这种方法的有效性和实用性。对于设计新的实验以确定转录因子的直接功能靶标的研究人员, 我们建议实验设计至少包含两个用于 DEG 鉴定的比较条件, 并随后分析 LOA 和  $\beta$  相结合进行的比较中的 DEG 列表。

我们确定的 HIF-1 靶基因富含代谢相关因子, 我们试图通过代谢组学来验证这一点。代谢组学分析只提供每个样本中代谢状态的时间快照。真正的代谢流量分析需要跟踪标记的代谢物的实验, 而这在线虫身上是做不到的。新陈代谢流量的变化可以建模, 但这样的模型需要限制其实用性的假设。然而, 像这项工作中进行的那种代谢分析创造了可以通过实验来探索的可检验的假说。

我们观察到, HIF-1 直接促进 PCK-1 PEP 羧基激酶的表达, PCK-1 是糖异生的关键介质, 并通过该途径促进代谢通量 (图 4F)。糖异生提供合成谷胱甘肽所需的代谢物, 并产生对抗 ROS 和氧化应激所需的 NADPH 还原等效物。事实上, 我们发现 PCK-1 突变体在低氧应激中的生存能力与 HIF-1 突变体一样差, 并且这两个突变体都可以通过补充 PCK-1 的产物 PEP 来拯救。此外, 补充抗氧化剂可以挽救这些突变体的低氧生存, 而 HIF-1 的激活可以抵消导致氧化应激的试剂造成的损害。我们对人类表达谱数据的荟萃分析表明, PCK2 是 PCK-1 的同源基因, 是由 HIF1a 调控的一组新的上下文相关靶点的一部分 (图 5)。我们的结果强调了 HIF-1 不仅需要促进无氧能量的产生, 还需要通过糖异生来动员抗氧化剂防御。事实上, 特异性地阻断糖异生可能为治疗 HIF-1 阳性肿瘤提供了一种方法。

尽管低氧和糖异生之间的联系早在 52 年前就已经被提出, 但我们的发现不仅仅是表明 HIF-1 直接与糖异生的关键介质结合并调节其表达。相反,

我们的工作强调了糖异生-PPP 代谢链在抵消缺氧期间氧化应激中的作用，并表明恢复氧化还原稳态是 HIF-1/糖异生调控相互作用的一种被低估的生理结果，这对低氧应激的生存至关重要。这一发现不仅对与缺氧有关的疾病(如癌症、心脏病、中风、慢性阻塞性肺病、脑瘫、肺动脉高压、新冠肺炎等)有影响，而且有助于我们理解好氧生物的进化及其适应具有可变氧气供应的陆地和水生栖息地的能力。

## 4 方法

### 4.1 转基因和转基因动物的产生

线虫种群是从 N2 种衍生而来的，在所有实验中都分析了两性种。种来源于线虫遗传中心。除非另有说明，线虫是在接种在 NGM 平板上的 OP50 大肠杆菌上培养的。GFP 标记的 HIF-1 转基因是从转基因 Ome 项目中获得的基因组融合蛋白。经氯霉素、链霉素、营养素三重选择筛选出多克隆插入培养物，通过测序筛选出单个克隆进行转基因验证。该构建物被引入到 *hif-1(aa4)* 缺失突变体的胚系基因组中，然后利用微粒子轰击稳定地整合到基因组中。所获得的稳定整合品系 *odIs131* 通过选择产生 100% 非 UNC 后代的单个两性株系，然后进行四次异交，使其成为纯合子。随后，将 *unc-119(ed3); odIs131* 菌株与 *unc-119(ed3); egl-9(Sa307)*; HIF-1(Ia4) 杂交，产生用于 ChIP-seq 的下列菌株：OR3349 *unc-119(ed3); hif-1(ia4); odIs131[hif1::gfp, unc-119(+)]* 和 OR3350 *unc-119(ed3); egl-9(sa307); hif-1(aa4); odIs131[hif-1::gfp unc-119(+)]*。

含有编码 HIF-1 结合位点和转录起始点的基因组序列的荧光金星转录报告基因被用来验证一些直接靶标。为了定量表达，将 *P<sub>myo-2</sub>::mCherry* 转基因与 *pck-1::Venus* 或 *rhy-1::Venus* 转基因一起引入基因组中。由于 *myo-2* 的表达在 RNA-seq 数据集中没有变化，因此 Venus 的表达在咽部归一化为 mCherry 的表达。对于每个实验，至少检查了两条独立的线。为了获得 pPCK-1::Venus 转基因动物，扩增了 pCK-1 ATG 上游的 2816 bp，并将其克隆到 pPD95.77-mVenus 中的 Venus 编码序列的上游。利用 Q5 定点定向突变技术 (Life Technologies Ltd.)，删除  $\Delta$  峰 (齐达内上游 1084-912 bp) 和  $\Delta$ HRE(1019-1003 上游)，分别产生 *P<sub>pck-1(\Delta Peak)</sub>::Venus* 和 *P<sub>pck-1(\Delta HRE)</sub>::Venus*。将这三种质粒 (100 ng/ L) 以 100 ng/ L 注入野生型动物体内，同时注射 *P<sub>myo-2</sub>::mCherry*(50 ng/ L) 共注射标记物。

为了获得  $P_{rhy-1}::Venus$  转基因动物, 扩增并克隆了如上所述的 Venus 编码序列上游的 1510 bp 的  $rhy-1$  ATG。Q5-定点突变 (Life Technologies Ltd)。用于去除启动子的  $\Delta$  峰 (上游 1170-617 bp)、 $\Delta HRE1$ (上游 971-956 bp)、 $\Delta HRE2$ (上游 946-931 bp) 和  $min(591p)$  变异体。每一质粒以 100 ng/ L 和  $P_{myo-2}::mCherry$ (50 ng/ L) 共注射标记物注入野生型动物体内。荧光分析如上。

## 4.2 荧光显微镜与图像分析

用 10 mM 四咪唑固定在 2% 琼脂糖垫上, 可在线虫体内观察到荧光蛋白。所有动物通过碱性漂白同步, 并在 L4 阶段可视化。使用 AxioImager M1M(Carl Zeiss, Thornwood, NY) 观察到含有  $odIs131[hif-1::gfp]$ 、 $PPCK-1::Venus$ 、 $Prhy-1::Venus$ 、 $n d/or r Pmyo-2::mCherry$  转基因动物的荧光图像。使用 5X(NA 0.15)、10X(NA 0.30) 或 40X(NA 1.3)PlanApo 物镜来检测荧光。用 ORCA 电荷耦合设备照相机 (Hamamatsu, Bridgewater, NJ) 通过使用 iVision 软件 V4.1(Biovision Technologies, Uwhclan, PA) 获取图像。选择曝光时间以捕捉所有样品的荧光强度动态范围的至少 95%。通过使用透射光图像获得线虫的轮廓来进行量化。使用斐济/ImageJ 2.1.0/1.53c55 计算每个轮廓内的平均荧光强度 (使用滚球滤光器减去背景盖板荧光后)。然而, 选择了一条典型的线进行定量, 分析了来自 20 到 100 只动物 (通常是 50 只) 的荧光强度, 并从两个生物重复中收集了它们。平均值代表动物的 Venus/mCherry 比率, 各个比率值归一化为每次实验中分析的对照动物的平均值。所有正态分布的数据用 GraphPad Prism 9.3.0 进行分析, 大多数情况下使用 ANOVA 和 Dunnett 的后检验校正进行多重比较。

## 4.3 定量 RT-PCR 检测基因表达

通过漂白两个装满妊娠动物的 60 mm 平板, 使相关基因型的线虫年龄同步。在 M9 中收集 L4 动物并清洗两次, 然后使用 TRIzol 试剂 (Ambion Life Technologies, Ref.15596026) 和 DirectZolRNAMiniPrep Plus(ZYMO, Cat. R2070), 均按制造商说明。在 Zymo 方案中, 用 DNA 酶处理去除残留的 DNA。分子生物学水 (Millipore, Cat.H2OMB0106) 洗脱 mRNA。用分光光度计 (NanoDrop 或 TecanPro) 测量核酸浓度, 并在 QRT 反应准备之前用水稀释到相同浓度。

为了检测 *odIs131[hif-1::gfp]* 基因的表达, 用两组引物检测内源性 *hif-1* 和 *odIs131[hif-1::gfp]* 转基因的表达水平。第一组扩增了所有 *hif-1* mRNA 转录本 3' 端的序列。第二组扩增了 *hif-1(ia4)* 缺失的序列。平均表达通常归一化为未经处理的野生型, 并来自两个或更多独立的生物试验。

为了评估缺氧暴露后 HIF-1 靶基因的表达, 野生型 (N2) 和 HIF-1(Ia4) 菌株在实验前被置于 0h(常氧) 和 4 h(缺氧) 的低氧处理中, 使用预先平衡到 0.5%O<sub>2</sub> 的氮置换缺氧室 (BioSpherix) 至少 2 h。使用 iTaq Universal SYBR®Green 一步试剂盒 (BioRad, Cat.1725150) 和 CFX Opus 384 实时聚合酶链式反应系统 (BioRad), 对每个基因型和条件下的四个生物复制进行分析, 每次反应 40 ng RNA。除非另有说明, 以肌动蛋白 (ACT-1) 作为归一化对照, 使用 *Delta CQ* 方法对结果进行分析。样本总是在同一平板上进行肌动蛋白对照试验。在每个平板上, 反应以重复进行, 并进行 1-3 个独立的平板技术复制。同一平板内 CQ 值之间标准偏差 >0.399 的技术重复在分析中被丢弃。使用棱镜 (Graphpad, 版本 9.3.0) 进行统计测试和数据表示, 具体测试细节可在相应的图例中找到。

在补充数据 6 中列出了引物。它们在 30 nm 处被用来扩增所指示的兴趣基因。

#### 4.4 产卵和卵滞留的特征

用碱性漂白法同步, 并在 M9 缓冲液中停留在 L1 期过夜。在胚胎发育到 L4 期后 43-46 h, 用解剖显微镜对成年动物体内的胚胎进行计数, 检测同步化的基因分型。分析了 40 到 60 只动物 (通常是 50 只), 并从三个生物副本中收集了它们。平均值代表每只动物未产的卵。

#### 4.5 ChIP-seq

对菌株 OR3349 和 OR3350 的 L4 期线虫, 按照它们的标准协议 42, 由 MODEM/modENCODE 联盟进行 ChIP-seq。通过漂白和 L1 停滞实现发育同步化。将滞留的 L1 细胞接种在种植有 OP50 细菌的 NGM 平板上, 在 20°C 下生长到 L4 阶段, 然后通过离心法收集它们。随后, 将颗粒状线虫在液氮中闪速冷冻, 并将其储存在 -80°C 下。将颗粒在冰上解冻, 加入 750 mL FA 缓冲液 (Roche Cat#11697498001 完全蛋白酶抑制剂 Cocktail 平板, 125 mL 10 mMMSF, 每 25 mL FA 缓冲液 nd 25 L1MDTT), 然后

将样品转移到 2mLKontes dust(Kimble Chase, Vineland, NJ) 中。样品在冰上浸泡 15 次, 用小的“A”槌浸泡两个周期, 每个周期之间保持 1 分钟。然后用大的“B”槌将样品浸泡 15 次, 共四轮, 每轮之间保持 1 分钟。样品用 2% 甲醛在室温下交联 30 min, 然后用 1M Tris pH 7.5 淬火。然后对样品进行超声波处理, 以将染色质剪切成 200-800 个核苷酸的 DNA 片段。

对于每个样本, 使用 15 g 抗绿色荧光蛋白抗体 (托尼·海曼和凯文·怀特赠送) 免疫沉淀 4 mg 蛋白质裂解物。对表达 AMA-1::GFP 的转基因动物的免疫沉淀物进行免疫印迹实验, 验证了多克隆山羊抗 GFP 抗体的有效性。对使用小鼠和山羊免疫球蛋白控制免疫沉淀进行了比较, 并与输入进行了比较以建立背景。通过比较 GFP 标记的 AMA-1(POL II) 和另一种直接针对天然 AMA-1 的抗体沉淀的天然蛋白的 ChIP-seq 数据集, 确定了与天然图谱的相似性, 得出两个生物复制的相关性分别为 0.93 和 0.95。GAMMABind G 琼脂糖珠 (GE Healthcare Life Sciences) 预洗, 在结合缓冲液和 0.1 mg/mL BSA 中封闭。样品在 4°C 下加入 100 L 的 50/50 微珠/缓冲液预净化, 然后离心。每个重复的样本被取出并汇集在一起作为总的染色质输入。对于每个重复, 将 15 g 抗体添加到 4°C 下 1:225 的稀释液中过夜, 然后与珠子混合物孵育另一个过夜。免疫沉淀物 (IP) 用冷裂解缓冲液清洗 4 次, 用冷 TE 清洗 2 次。将小球重新悬浮在洗脱缓冲液 (10 mM EDTA, 1% 十二烷基硫酸钠, 50 mM 三氯化钠 pH 8) 中, 在 65°C 孵育 10 分钟。样品被离心, 上清液被转移到新的试管中。将微球重新悬浮在 29%TE 和 0.67%SDS 中, 并立即离心。洗脱上清液混合, 在 65°C 温和摇动过夜孵育。将染色质输入样品在 60°C 下孵育一夜, 然后将蛋白酶 K 和十二烷基硫酸钠分别加入到最终浓度为 0.1 mg/mL 和 0.01% 的浓度中。第二天, 将输入材料在 70°C 孵化 20 分钟。在每个 IP 中加入蛋白酶 K, 在 50°C 孵育 2 h, 在染色质输入中加入 0.017 mg/ml 的 RNaseA, 在 37°C 孵育 2 h, 用 MinElute 柱 (QIAGEN, Valencia, CA) 纯化 DNA, 在 13 L (MinElute 试剂盒提供的洗脱缓冲液) 中洗脱。纯化后的进样再加入 48 L EB。样品在 -20°C 的温度下保存。

两个生物副本的富含 DNA 片段和输入对照 (来自同一样本的基因组 DNA) 用于文库制备和测序。将样品转换成文库, 并使用 Ovation 超低 DR Multiplex Systems (NuGen Technologies, San Carlos, CA) 按照制造商的方案进行多路复用, 只是使用 QIAGEN MinElute PCR 纯化试剂盒分离 DNA。简而言之, 用 1 L 的输入 DNA 和 10 L 的 IP DNA 用 NuGen 超

低文库试剂盒制备测序文库。样品是根据制造商的协议准备的。对 Illumina HiSequation 2000/2500/4000 进行测序, 得到用于输入的 6.5-14.1 M 单端读数的范围, 以及 OR3349 和 OR3350 的两个重复。所有样本的 Phred 得分从 32 到 38 到至少 47 个碱基。

使用 Burrow-Wheeler 比对 (BWA) 将 Illumina 测序数据与参考基因组进行比对。数据与基因组版本 WS245 比对。40.4-85% 的读数保持一致。通过用于 MODEM/MODENCODE42, 57 的 SPP 芯片序列处理流水线标准来调用在对齐读数中显著丰富的峰区。大于 0.1% 的不可再生性发现率 (IDR) 的峰被用来产生最终的峰集。OR3349 和 OR3350 的最低富集值分别为 93.8 和 50.8。可在此处获得现代/编码方法的大致轮廓: <https://www.encodeproject.org/Pipeline/ENCPL631XPY/>。分析工具可在 GitHub 上获得: <https://github.com/ENCODE-DCC/chip-seqpipeline2/releases/tag/v1.3.5.1>。

#### 4.6 RNA-seq

用次氯酸盐处理妊娠成虫和在 M9 中孵化 15-17h 的胚胎获得发育同步化动物。饥饿抑制的 L1 被镀在 NGM 板上, 并在 20°C 下生长到 L4 阶段。使用 Trizol(Invitgen) 结合滴打碎器裂解从动物身上提取总 RNA, 每个基因有四个生物重复。使用带 PolyA 选择的 Illumina Truseq(Genewiz 或 RUCDR) 为每个样品/复制准备一个 mRNA 文库 (单端, 50bp 读取)。在快速运行模式 (RUCDR) 下, 在 Illumina HiSeq2000(GeneWiz) 或 Illumina HiSeq 2500 上跨两个通道对文库进行测序, 得到每个样本 38-58M 单端读取的范围, 峰值平均 Phred 得分为 39。读数被映射到线虫基因组 (WS245) 和 STAR 2.5.1a 产生的基因计数。88%-91% 的读数是一致的。使用广义线性模型功能和标记离散度估计, 使用 Edger 对基因计数进行归一化和统计分析。多维尺度分析显示, 在四个生物复制中紧密聚集, 其中 HIF-1 活跃和不活跃的基因型之间有明显的分离。似然比检验是以本杰米尼和霍奇伯格校正的方式在基因型间成对进行的。比较了 L4 期动物在有氧条件下 HIF-1 活性的基因型, 包括 *egl-9(sa307)* 突变体和 *egl-9(sa307); hif-1(ia4); odIs131[hif-1::gfp]* 突变体; 以及 HIF-1 失活的基因型, 包括野生型和 *egl-9(sa307); HIF1(Ia4)* 双突变体。在以下四项配对比较中, 如果基因在同一方向 (向上或向下) 差异表达, FDR<0.01, 则被认为是 HIF-1 依赖的基因: (1)*egl-9* 与 N2; (2)*egl-9* 与 *egl-9 hif-1*; (3)*odIs131; EGL-9* 与



N2; (4)*odIs131;egl-9* 与 *egl-9 hif-1*。RNA-seq 数据集可在 NIH/NCBI GEO 上通过登录号 GSE173581 获得。

#### 4.7 用 BETA basic 和 LOA 识别直接靶标

分析了两个成对的 RNA-seq 比较 (野生型与 *egl-9* 单突变体, 以及由 *odIs131* 转基因挽救的野生型与 *egl-9 hif-1* 双突变体), 测量了两种比较中相同基因的差异表达。采取有条件的方法, 检查第一次比较 (野生型与 *egl-9* 单一突变体) 中的信息, 看看它是否影响第二次比较 (通过转基因挽救的野生型与 *egl-9 hif-1* 双突变体) 的解释。利用 Luperchio 等人的方法, 根据第一次比较的结果, 第二次比较中的基因被分成两组, 一组包括在第一次比较中发现的差异表达的基因, 第二组包括发现没有差异表达的基因。为了估计哪些基因是差异调控的, 使用 0.01 的 FDR 在两次比较之间生成一个重叠列表, 称为 LoA Combo 1。使用相同的条件方法来询问来自单独实验 (*egl-9 hif-1* 双突变体与 *egl-9* 单一突变体) 的比较信息是否影响第二次比较 (*egl-9 hif-1* 双突变体与通过 *odIs131* 转基因挽救的 *egl-9 hif-1* 双突变体) 的解释。这两个比较之间的重叠列表被称为 LoA Combo 2。Beta Basic 用于识别线虫基因组 41 的 WS245 注释的 *egl-9* 的潜在直接靶点。使用了以下参数: TSS 的 15 kb, FDR 校正为 0.01, 单尾 KS 检验截止值为 0.01。输入文件由 IDR 阈值峰值和 LOA 组合 1 和组合 2 之间重叠 (交集) 的差异表达式 Log2FC 和 FDR 值的 .ed 文件组成。用这种方法鉴定的 HIF-1 峰/基因组合被标记为 “LoA/beta”。分析工具可在 GitHub 网站上获得: <https://github.com/shahlab/homy-multiomics>。

在单独的分析中, 获得了四个成对的 RNA-SEQ 比较的重叠, 而没有 LOA 分析的好处 (即, 满足每个列表的 FDR 截止值 0.01 的四个基因列表的简单交集)。如上所述, 重叠列表用于 BETA 分析, 以确定 HIF-1 的潜在直接靶点。用这种方法鉴定的 HIF-1 峰/基因组合被标记为 “仅 BETA”。

#### 4.8 Motif 识别和富集

从 UCSC 基因组浏览器中提取每个 ChIP-Seq 峰 ( $\pm 100$  bp, 重复序列用 N 掩蔽) 中间的序列, 并将其输入 Meme-Chip Tool v5.4.1, 网址为 [memesuite.org](http://memesuite.org)。使用以下参数: JASPA2018 核心脊椎动物数据库 (非冗余), 模体宽度从 6 到 15 个碱基对, 一阶背景模型, 链截止值 0.05。

## 4.9 HIF-1 结合位点转录因子占位的鉴定

从 modENCODE 文库获得在 L4 阶段进行 CHIP-SEQ 的所有转录因子的结合位点坐标 (IDR 阈值峰, FDR<0.01)。为了确定这些峰与 OR3350 的 HIF-1 芯片序列峰重叠的实例数, 我们开发了一个 R 脚本, 可从 <https://github.com/shahlab/hypoxia-multiomics> 获得。

## 4.10 代谢组学

利用液质联用 (LC-MS) 技术对 L4 期野生型和 EGL-9(Sa307) 突变株在好氧条件下的代谢谱进行了研究, 共检测了 9 个独立生物重复的 558 个代谢物。用碱性漂白的方法使所有的基因型年龄同步, 并将滞留的 L1 幼虫接种在 100 mm 的平板上, 平板上装有标准的 NGM 培养液和 OP50。动物在 20°C 下生长到 L4 阶段, 此时它们被洗入 50 毫升的圆锥形底管中, 并停留 15 分钟。动物用 3×40 毫升无菌 M9 洗涤, 以 2000 转/分的速度离心 5 min。最后的颗粒被转移到 15 mL 的圆锥形底管中, 在液氮中闪速冷冻至 80°C, 直到送到代谢素 (–) 进行代谢组处理。样品用甲醇提取, 每个萃取物被分成四个组分: 两个分别用于正离子模式电喷雾电离 (ESI) 的反相 (RP)/UPLC-MS/MS 法分析, 一个用于 RP/UPLC-MS/MS 负离子模式电喷雾电离 (ESI) 分析, 一个用于 HILIC/UPLC-MS/MS 负离子模式电喷雾电离 (ESI) 分析。使用定制软件 (新陈代谢) 提取原始数据并进行峰值识别, 然后与基于认证标准的文库进行比较。使用 ArrayStudio 中的 R 进行统计分析 (Welch ‘s t 检验, 多次比较校正)。

## 4.11 百草枯存活试验

根据 Senchuk 等人 59 中描述的急性百草枯敏感性试验进行百草枯敏感性测定。将动物置于 20°C 的 L4 阶段, 每小时统计一次存活时间, 直到所有动物死亡。每个试验的每个品系有 20 到 100 只动物 (通常是 50 只) 用于检测, 平均存活 15 小时的动物百分比来自五个独立的生物试验。

## 4.12 低氧存活试验

*pck-1* 缺失突变体在有氧条件下是可行的, 很可能是因为另外两个 PEP 羧基酶, *pck-2* 和 *pck-3*, 补偿了基线功能。*pck-2* 和 *pck-3* 均不具有 HIF-1 依赖的调节作用, 提示 *pck-1* 是 HIF-1 诱导的特异性异构体。

对于低氧饥饿，M9 液与 0.5% $O_2$  预平衡至少 1 小时。线虫在标准的 NGM 培养皿上与食物一起生长。取同龄 L4 动物 (n 20)，用 M9 缓冲液清洗，然后在 20°C 的盖管中用低氧 M9 液孵育指定时间。FUDR(5-氟-2-脱氧尿嘧啶核苷，VWR 102573-230VWR) 在最终浓度为 50  $\mu$ M 时加入，以防止子代的产生，从而干扰分析。低氧饥饿后，用离心机收集线虫，放入常氧下的标准 NGM 平板中，24 小时后恢复。通过评估动物的活动和对铂金钢丝钩轻轻触摸的反应来计算动物的存活率。每个试验使用了 20 到 100 只动物 (通常是 50 只) 进行检测，平均 48 小时存活动物的百分比来自四个独立的生物试验。另一组单独的对照 (常氧) 动物以同样的方式进行分析，但没有暴露在低氧环境中。

对于胚胎在低氧条件下的存活，我们使用了以前发表的方案 29。将年龄同步化线虫与食物一起种植在标准的 NGM 培养板上，直到 L4 期后 48h，用次氯酸盐处理裂解线虫，释放其宫内胚胎。将胚胎放入新的 NGM 培养皿中，计数后置于预先平衡的 (0.5%  $O_2$ ) 低氧舱中，在 25°C 下培养 24 h。缺氧处理后，取出培养皿，在 25°C 常氧条件下继续培养 24 h，计算未孵化胚胎数。每个试验使用了 50 到 100 个胚胎进行分析，平均有百分比的胚胎存活来自六个独立的生物试验。另一组单独的对照 (常氧) 动物以同样的方式进行分析，但没有暴露在低氧环境中