

# NGHIÊN CỨU SẢN XUẤT MỘT SỐ ENZYME PHÂN HỦY LIGNOCELLULOSE

**GS.TS Trương Nam Hải**

*Viện Công nghệ Sinh học, Viện Hàn lâm Khoa học và Công nghệ Việt Nam*

Các nhà khoa học thuộc Viện Công nghệ Sinh học, Viện Hàn lâm Khoa học và Công nghệ (KH&CN) Việt Nam đã thực hiện thành công nhiệm vụ hợp tác quốc tế về KH&CN theo Nghị định thư với CHLB Đức: “Nghiên cứu sản xuất một số enzyme phân hủy lignocellulose trên cơ sở khai thác dữ liệu metagenome”, mã số NĐT.50.GER/18. Thông qua đó đã xây dựng được các quy trình công nghệ tiên tiến, cho phép khai thác và nghiên cứu tính chất của các enzyme phân hủy lignocellulose từ dữ liệu DNA đa hệ gen vi khuẩn bằng các công cụ tin sinh và công nghệ gen.

## Tìm kiếm nguồn enzyme mới

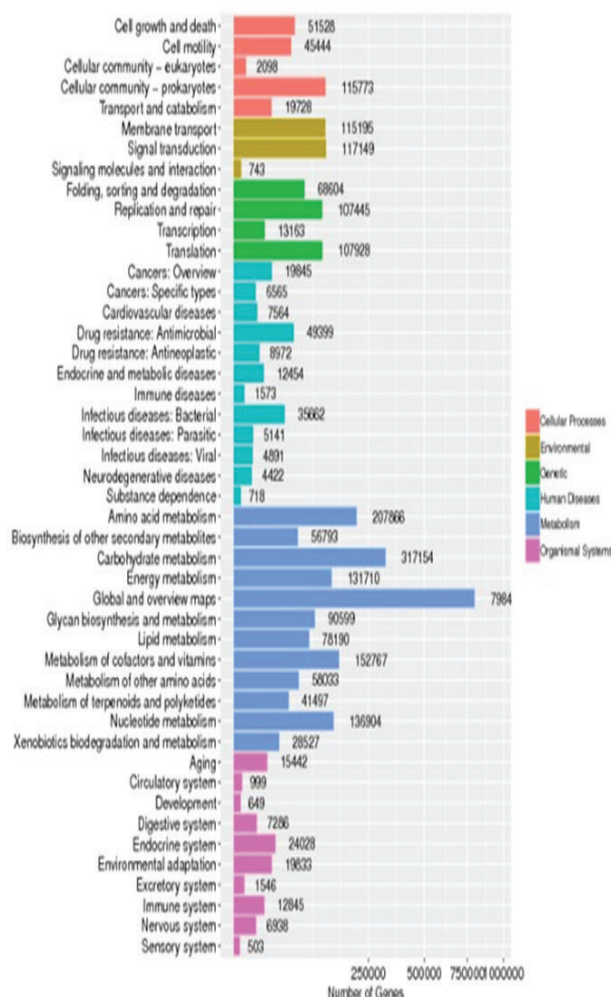
Lignocellulose từ phế phụ phẩm nông - lâm - công nghiệp là nguồn sinh khối tái tạo vô cùng dồi dào trên Trái đất. Đây là nguồn nguyên liệu quan trọng để phát triển nền kinh tế sinh học. Là quốc gia có thế mạnh về nông nghiệp, tổng nguồn sinh khối hàng năm của Việt Nam vào khoảng 119 triệu tấn, trong đó có khoảng 40 triệu tấn rơm rạ, 8 triệu tấn trấu, 6 triệu tấn bã mía và trên 50 triệu tấn vỏ cà phê/vỏ đậu/phế thải gỗ... Tuy nhiên, chỉ một phần rất nhỏ lượng sinh khối này được sử dụng để trồng nấm, đun nấu và phát triển điện sinh khối. Còn lại hầu hết bị vứt bỏ, gây ô nhiễm môi trường và lãng phí nguồn nguyên liệu sẵn có. Trong khi đó, từ nguồn nguyên liệu này có thể sản xuất ra nhiều sản phẩm có giá trị kinh tế cao sau khi được chế biến bằng các quá trình sinh học và hóa học.

Trong tự nhiên, lignocellulose được chuyển hóa sinh học bởi các enzyme thành các đường đơn và được vi khuẩn sử dụng như nguồn thức ăn. Nguồn enzyme này rất đa dạng về hoạt tính và tính chất sinh - hóa phụ thuộc vào môi trường sống của vi khuẩn. Do đó, việc nghiên cứu nhằm khai thác các enzyme này từ vi khuẩn là rất cần thiết. Mặc dù vậy, do hạn chế về mặt kỹ thuật nên hiện nay chỉ có khoảng 1% vi khuẩn có thể nuôi cấy trong phòng thí nghiệm. Điều này hạn chế khả năng nghiên cứu tính chất của các enzyme từ nguồn vi khuẩn trong tự nhiên. Để khắc phục khó khăn này, các nhà khoa học đã phát triển kỹ thuật đa hệ gen (Metagenomics) cho phép khai thác và nghiên cứu các enzyme phân hủy lignocellulose từ vi khuẩn không thông qua nuôi cấy.

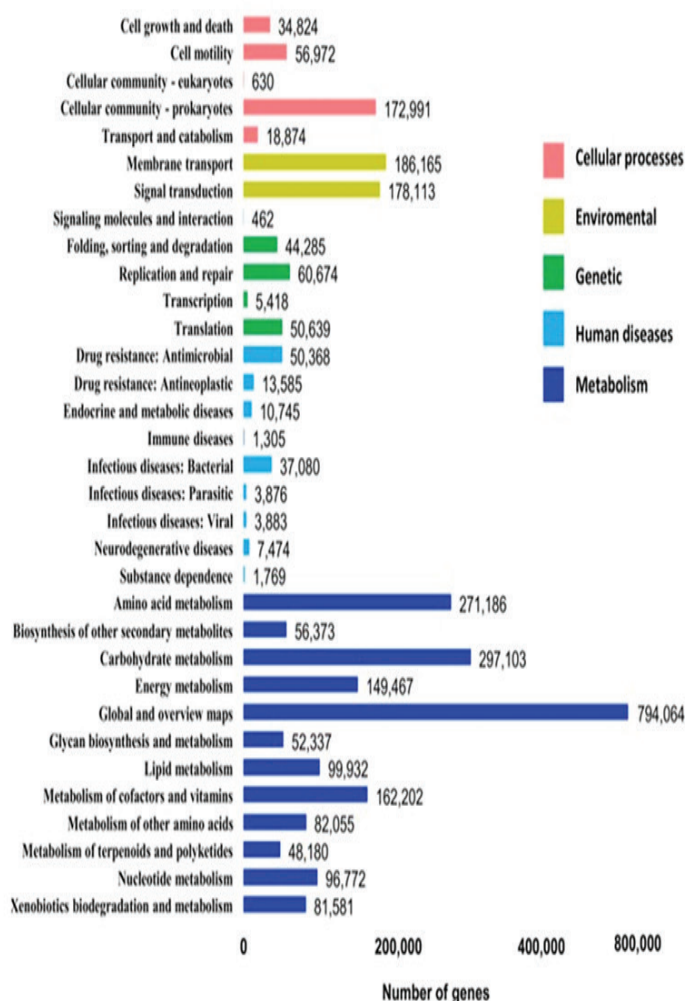
## Giải mã bằng metagenome

Nhằm khai thác hiệu quả các protein và enzyme tham gia quá trình chuyển hóa sinh học nguồn sinh khối lignocellulose từ 2 hệ sinh thái khác nhau là khu hệ vi khuẩn trong dạ cỏ dê (dê Cỏ, dê Bách Thảo sống chăn thả tại Ninh Bình và Thanh Hóa) và hệ vi khuẩn xung quanh khu nấm mục trắng trên thân gỗ đang phân hủy (thu tại Vườn quốc gia Cúc Phương), Viện Công nghệ Sinh học đã đề xuất và được Bộ KH&CN phê duyệt thực hiện đề tài: “Nghiên cứu sản xuất một số enzyme phân hủy lignocellulose trên cơ sở khai thác dữ liệu metagenome (ký hiệu: MetagenLig)”. Mục tiêu của đề tài là thiết lập được bộ dữ liệu các gen mã hóa enzyme/protein phân hủy lignocellulose trên cơ sở phân tích dữ liệu metagenome và sản xuất được một số enzyme phân hủy lignocellulose có hoạt tính cao.

(A)



(B)



Kết quả chú giải gen theo các nhóm chức năng của mẫu quần xã vi khuẩn dạ cỏ dê (A) và quanh nấm mục trắng (B).

Kết quả dự đoán các gen cho thấy, có 4.222.317 gen được khai thác từ cơ sở dữ liệu DNA đa hệ gen vi khuẩn của dạ cỏ dê và 3.884.879 gen được khai thác từ cơ sở dữ liệu đa hệ gen vi khuẩn quanh nấm mục trắng. Các gen ở vi khuẩn dạ cỏ dê được chia thành 6 nhóm gen chức năng liên quan tới 46 quá trình sinh học, còn ở vi khuẩn quanh nấm mục trắng là 5 nhóm gen chức năng liên quan tới 33 quá trình sinh học. Ngoài 5 nhóm chức năng giống nhau giữa

hai hệ sinh thái (trao đổi chất, các quá trình hoạt động của tế bào, di truyền, môi trường và các bệnh liên quan đến người), ở vi khuẩn dạ cỏ dê còn có nhóm gen chức năng liên quan đến các hệ thống sinh học của dê như các hệ thống: tuần hoàn, tiêu hóa, nội tiết, bài tiết, miễn dịch, thần kinh, cảm giác... và quá trình lão hóa. Vì vậy, vi khuẩn trong hệ tiêu hóa được ví như bộ gen thứ hai liên quan tới các quá trình sinh học của vật chủ.

Trong tổng số 4.222.317 gen được chú giải chức năng từ dữ liệu DNA đa hệ gen của vi khuẩn dạ cỏ dê, có 67.540 gen tham gia quá trình phân hủy sinh học lignocellulose. Con số này ở vi khuẩn quanh nấm mục trắng là 22.226/3.884.879 gen (bảng 1).

Kết quả khai thác gen chuyển hóa lignocellulose từ dữ liệu DNA đa hệ gen vi khuẩn dạ cỏ dê cho thấy, vai trò rất quan trọng của

**Bảng 1. Kết quả chú giải chức năng protein, enzyme tham gia quá trình chuyển hoá sinh học lignocellulose của mẫu vi khuẩn dạ cỏ dê và vi khuẩn quanh nấm mục trắng.**

| Số lượng gen được chú giải tham gia các quá trình phân hủy lignocellulose | Quần xã vi khuẩn từ dạ cỏ dê | Quần xã vi khuẩn quanh nấm mục trắng |
|---------------------------------------------------------------------------|------------------------------|--------------------------------------|
| Số lượng protein và enzyme tham gia quá trình tiền xử lý lignocellulose   | 2.670                        | 907                                  |
| Số lượng enzyme tham gia phân hủy cellulose                               | 21.029                       | 8.301                                |
| Số lượng enzyme tham gia phân hủy hemicellulose                           | 43.841                       | 13.018                               |
| Tổng số                                                                   | 67.540                       | 22.226                               |

chi vi khuẩn *Prevotella* trong việc tăng cường chuyển hóa thức ăn lignocellulose của dê. Đây là kết quả lần đầu tiên được phát hiện ở dê và nó là bằng chứng để nhận diện vai trò chỉ thị của nhóm vi khuẩn này trong phân tích quần xã vi sinh vật.

Từ 89.766 gen tham gia quá trình phân hủy sinh học lignocellulose nêu trên, đề tài đã chọn 30 gen để nghiên cứu tính chất và đã khẳng định được 9 enzyme có hoạt tính tốt trong đó có các enzyme quan trọng như endo-glucanase, xylan beta-xylosidase, beta-glucosidase và LPMO. Các enzyme được lựa chọn có cấu trúc rất đặc biệt với hoạt tính tốt. Trình tự axit amin của các enzyme này có mức độ khác biệt cao so với các enzyme đã nghiên cứu. Từ 9 enzyme đã nghiên cứu có 2 enzyme là endo-glucanase EG5 (có cấu trúc vùng chức năng BM72-CBM72) và beta-glucosidase (có trình tự axit amin và cấu trúc mới GH3-FN3-GH31 so với các enzyme đã nghiên cứu) đã được lựa chọn để bước đầu xây dựng quy trình sản xuất trong nổi lên men.

### Nhiệm vụ mang lại

Có thể kể đến những kết quả chính của đề tài như sau: (1) Xây dựng được 2 bộ dữ liệu DNA chứa hàng triệu gen mã hóa cho các protein và enzyme đóng vai trò quan trọng đối với các quá trình sinh học của vi khuẩn; khai thác được một phần nhỏ các gen của hai bộ dữ liệu liên quan đến quá trình trao đổi chất lignocellulose; (2) Xây dựng được các quy trình khai thác hiệu quả gen từ cơ sở dữ liệu DNA và các quy trình công nghệ sàng lọc, biểu hiện và tinh chế các protein, enzyme khai thác được từ bộ dữ liệu gen; (3) Đánh giá được sự đa dạng quần xã vi khuẩn trong hai hệ sinh thái đặc thù (dạ cỏ dê và nấm mục trắng trên thân gỗ đang phân hủy trong rừng Cúc Phương).

Một trong những thách thức lớn khi ứng dụng kỹ thuật metagenomics là làm thế nào để có thể nhận dạng nhanh và chính xác các gen quan tâm từ dữ liệu DNA với hàng triệu gen giả định. Trong quá trình thực hiện đề tài, 2 đối tác là GS.TS Juergen Pleiss (Đại học Stuttgart) và GS.TS Wolfgang Streit (Đại học

Hamburg) của CHLB Đức đã chuyển giao và đào tạo cho cán bộ của Viện 2 quy trình công nghệ rất quan trọng. Một là công nghệ chú giải gen (gene annotation) bằng trí thông minh nhân tạo, giúp chú giải các gen quan tâm dựa trên ứng dụng mô hình đại diện của các nhóm protein và enzyme đã được nghiên cứu cấu trúc và tính chất sinh hóa. Phương pháp này cho phép nhận dạng nhanh và chính xác các gen đích từ dữ liệu DNA. Các gen được nhận dạng sau đó sẽ được biểu hiện và nghiên cứu tính chất sinh học. Hai là, công nghệ biểu hiện vô bào, cho phép sàng lọc nhanh với số lượng lớn các gen dựa trên hệ biểu hiện vô bào. Hai quy trình công nghệ này đã giúp cho nhóm nghiên cứu khai thác nhanh và hiệu quả các gen quan tâm từ dữ liệu DNA đa hệ gen.

Bên cạnh đó, đề tài đã đào tạo được 2 nghiên cứu sinh, công bố 10 bài báo trên các tạp chí khoa học trong nước và quốc tế.