

TIÊU CHUẨN QUỐC GIA

TCVN xxxxx-3:2025

Xuất bản lần 1

GÕ – XÁC ĐỊNH LOÀI BẰNG CÔNG NGHỆ QUANG PHỔ KHỐI LƯỢNG – PHẦN 3: PHƯƠNG PHÁP XÂY DỰNG CƠ SỞ DỮ LIỆU

Wood – Method of wood identification using DART TOFMS technology – Part 3: Database development method

Dự thảo lần 1

HÀ NỘI – 2024

Lời nói đầu

TCVN xxxxx-3:2025 do Viện Khoa học Lâm nghiệp Việt Nam biên soạn, Bộ Nông nghiệp và Phát triển Nông thôn đề nghị, Tổng cục Tiêu chuẩn Đo lường Chất lượng thẩm định, Bộ Khoa học và Công nghệ công bố.

TIÊU CHUẨN QUỐC GIA

Gỗ - Xác định loài bằng công nghệ quang phổ khối lượng – Phần 3: Phương pháp xây dựng cơ sở dữ liệu

Wood – Method of wood identification using DART TOFMS technology – Part 3: Database development method

1 Phạm vi áp dụng

Tiêu chuẩn này quy định phương pháp xây dựng cơ sở dữ liệu phục vụ công tác giám định gỗ bằng công nghệ quang phổ khối lượng (DART-TOFMS).

2 Tài liệu viện dẫn

Các tài liệu viện dẫn sau đây là cần thiết để áp dụng tiêu chuẩn này. Đối với các tài liệu viện dẫn ghi năm công bố thì áp dụng bản được nêu. Đối với các tài liệu viện dẫn không ghi năm công bố thì áp dụng phiên bản mới nhất, bao gồm cả các sửa đổi, bổ sung (nếu có).

Schmitz, N., Beeckman, H., Blanc-Jolivet, C., Boeschoten, L., Braga, J. W., Cabezas, J. A., ... & Zuidema, P. A. (2020). Tổng quan về các phương pháp sử dụng trong giám định gỗ. Hướng dẫn về các phương pháp giám định gỗ.

TCVN 14121: 2024 Gỗ - Phương pháp định loại dựa vào đặc điểm cấu tạo thô đại và hiển vi

TCVN xxxxx-2: 2025 Gỗ - Xác định loài bằng công nghệ quang phổ khối lượng – Phần 2: Phương pháp lấy mẫu.

3 Thuật ngữ và định nghĩa

Trong tiêu chuẩn này, áp dụng các thuật ngữ và định nghĩa được nêu trong TCVN xxxxx-1: 2025, và TCVN 14121: 2024.

4 Nguyên tắc chung

Cơ sở dữ liệu là một trong những yếu tố quan trọng nhất đối với việc giám định gỗ bằng công nghệ quang phổ khối lượng (DART-TOFMS). Tiêu chuẩn này mô tả quy trình xây dựng cơ sở dữ liệu để tích hợp với cơ sở dữ liệu ForeST[©].

CHÚ THÍCH: ForeST[©] - Forensic Spectra of Trees (hay còn gọi là cơ sở dữ liệu quang phổ pháp y của cây) được xây dựng bởi Trung tâm giám định gỗ Hoa Kỳ (US-WISC) tại Ashland, OR, Hoa Kỳ. Đây là cơ sở dữ liệu lớn nhất về quang phổ khối lượng được thu thập từ các loại gỗ trên thế giới, bao gồm hơn 20.000 phổ khối từ 16.000 loài khác nhau. Phần lớn các loài có trong danh mục CITES và các loại gỗ thương mại phổ biến hiện nay để phục vụ giám định gỗ bằng công nghệ DART-TOFMS.

Trong quá trình xây dựng cơ sở dữ liệu phải tuân thủ các quy định về cách đặt tên, tạo tệp dữ liệu và thư viện NIST để đồng nhất và tích hợp được với cơ sở dữ liệu ForeST[©].

US-WISC khuyến nghị ForeST[©] nên được sử dụng với hệ thống JEOL Accu-TOF MS. Do tính phức tạp về mặt hóa học của gỗ nên việc phân tích để phân loại gỗ thông qua quang phổ khối lượng có thể sử dụng, điều chỉnh và hiệu chuẩn nhiều thông số kỹ thuật ở thiết bị nhằm nhất quán trong việc xây dựng cơ sở dữ liệu.

5 Thiết bị, dụng cụ và hóa chất

5.1 Thiết bị JEOL Accu-TOF MS, đây là hệ thống thiết bị thu thập phổ khối (mặc dù không phải là duy nhất được sử dụng để giám định gỗ). Tuy nhiên, đây là máy quang phổ khối phát triển nhất với cơ sở dữ liệu và phần mềm đi kèm tiên tiến nhất.

5.2 Phần mềm Mass Mountaineer, đây là phần mềm chuyên dụng cho công nghệ DART TOFMS bao gồm tích hợp tất cả các công cụ xử lý thống kê tiêu chuẩn. Phần mềm này có nhiều tùy chỉnh phục vụ các nghiên cứu và phân tích khác. Khuyến nghị sử dụng trong khoa học pháp y.

5.2 Tủ sấy, đối lưu cưỡng bức có khả năng luôn duy trì nhiệt độ ở (103 ± 2) ⁰C trong quá trình sấy mẫu đến khi đạt được khối lượng không đổi. Hơi ẩm sẽ được thoát ra ngoài qua lỗ thông.

5.3 Bình hút ẩm, kín có chứa chất hút ẩm (ví dụ: silica gel, canxi clorua, v.v...) để duy trì môi trường khô.

5.4 Dao cắt mẫu gỗ, sử dụng để trích mảnh gỗ mỏng từ mẫu gỗ.

5.5 Nhíp gắp mảnh gỗ mỏng, sử dụng để gắp mảnh gỗ mỏng khi thao tác thu thập phổ khối.

5.6 Polyetylen glycol 600 (PEG 600), một polyme có trọng lượng phân tử trung bình là 600 được sử dụng để hiệu chuẩn thiết bị trước khi thực hiện việc thu thập quang phổ khối lượng.

5.7 Methanol, đạt chuẩn HPLC (sắc ký lỏng hiệu năng cao) được sử dụng để vệ sinh thiết bị.

5.8 Khí Nitơ và Heli, độ tinh khiết 99,99%.

6 Chuẩn bị mẫu để xây dựng cơ sở dữ liệu

6.1 Lấy mẫu và chuẩn bị mẫu theo TCVN xxxxx-2: 2025.

6.2 Các mẫu gỗ phải được làm sạch bề mặt, yêu cầu không được xử lý biến tính bằng hóa chất, nhiệt độ, không được lẫn vecni, sơn, dầu, hoặc các chất phủ khác.

6.3 Thực hiện giám định khẳng định các mẫu gỗ lấy trên cây đứng (bằng khoan tăng trưởng) theo phương pháp hình thái học (tức sử dụng hoa, lá hoặc quả của cây đó để xác định). Đối với mẫu gỗ được thu thập trên cây đã chặt hạ, khúc gỗ tròn hay gỗ xẻ thì thực hiện giám định khẳng định theo TCVN 14121: 2024 và có thể sử dụng hoa, lá hoặc quả để xác định, kiểm chứng kết quả nếu nghi ngờ theo phương pháp hình thái học.

6.4 Sau khi giám định khẳng định, các mẫu được ký hiệu và mã hóa dữ liệu với đầy đủ thông tin về tên loài, chi, họ, nguồn gốc xuất xứ và một số thông tin khác,...

6.5 Sau khi chuẩn bị, bảo quản mẫu thử trong các điều kiện sao cho độ ẩm không thay đổi trước khi sử dụng.

7 Cách tiến hành xây dựng cơ sở dữ liệu

7.1 Thu thập quang phổ khối lượng (phổ khối)

Thu thập phổ khối được tiến hành theo hướng dẫn sử dụng của thiết bị. Các phổ khối được hiệu chuẩn bằng dung dịch tham chiếu có khối lượng đã biết (dung dịch PEG 600).

Để lưu và chia sẻ/tích hợp được cơ sở dữ liệu phổ khối của các mẫu gỗ tham chiếu, điều quan trọng là phải cung cấp được các thông tin chi tiết và đặt tên tệp cho từng phổ khối riêng của các mẫu gỗ tham chiếu đó.

7.2 Chuẩn bị và lọc cơ sở dữ liệu

Trước khi xây dựng mô hình thống kê, cần phải loại bỏ các giá trị (ở đây là phổ khối) ngoại lệ khỏi các tập huấn luyện. Hawkins (1980) mô tả 'giá trị ngoại lệ' là một giá trị khi quan sát thấy lệch quá nhiều so với các giá trị khác (tức khác biệt và không đồng nhất với các giá trị còn lại).

Sử dụng phần mềm Mass Mountaineer để xây dựng bản đồ nhiệt, cho phép so sánh đồng thời tất cả dữ liệu quang phổ và xác định các giá trị ngoại lệ để loại bỏ.

Kiểm tra chéo các quang phổ với các quang phổ đã có trong cơ sở dữ liệu và với các quang phổ của các chất gây ô nhiễm đã biết để kiểm chứng kết quả.

CHÚ THÍCH: Những lý do có thể dẫn đến giá trị ngoại lệ là: (1) Các mẫu được dán nhãn/ký hiệu sai hoặc xác định sai; (2) Việc thu thập phổ khối không được thực hiện đúng quy trình; (3) Lấy mẫu từ phần gỗ dác hoặc có dính phần gỗ dác; (4) Mẫu đã bị xử lý/nhiễm hóa chất, biến tính.

7.3 Xây dựng mô hình

Chỉ sử dụng phổ tham chiếu có sự đồng thuận (tức cường độ tương đối của các phân tử là giống/tương đồng nhau).

Đưa ra phổ tham chiếu để kiểm tra về hiện tượng đồng nhất về cường độ của các phổ và xác thực mô hình.

Xây dựng các mô hình thống kê để phân loại loài/nguồn gốc bằng cách sử dụng cùng số lượng khối phổ của mỗi loài/nguồn gốc và lựa chọn loại phân tích thống kê phù hợp (phụ thuộc vào loài/nguồn gốc (nhóm) đang được nghiên cứu). Sử dụng các phân tích thống kê như: (i) PCA (Principal Component Analysis), (ii) KDA (Kernel Discriminant Analysis) hoặc (iii) DAPC (Discriminant Analysis of Principal Components) với phần mềm Mass Mountaineer được khuyến nghị cho giám định pháp y vì phần mềm này dễ dàng tạo ra kết quả định lượng và cung cấp ước tính xác suất, thường được sử dụng tại tòa án để mô tả mức độ chính xác của các phân tích.

CHÚ THÍCH: Việc ứng dụng DART - TOFMS để xác định loại gỗ là một kỹ thuật tương đối mới (từ năm 2012), do đó các lựa chọn thuật toán để xây dựng mô hình không giới hạn ở các loại phân tích thống kê như trên vì còn nhiều các loại công cụ toán học khác có thể sử dụng.

7.4 Thẩm định và tối ưu hóa mô hình

Trước khi hoàn thiện mô hình, dữ liệu khối phổ nên được sàng lọc và thẩm định lại để tối ưu hóa độ chính xác trong phân loại. Tùy thuộc vào thuật toán phân loại, việc tối ưu hóa mô hình có thể được thực hiện như sau:

- Giảm số lượng biến (tổ hợp ion) bằng cách lựa chọn cẩn thận các thành phần trong PCA.

- Lựa chọn ion để xây dựng mô hình bằng cách sử dụng phân tích tỷ lệ Fisher hoặc kiểm tra bằng cách quan sát trực quan cẩn thận trên các bản đồ nhiệt.

- Dung sai khối lượng cho việc phân loại và cài đặt ngưỡng tương đối (tham số tiền xử lý) và số lượng ion (tham số phân loại) có thể được sàng lọc theo cách tự động khi sử dụng thuật toán Random Forest hoặc bất kỳ thuật toán nào phụ thuộc vào khung dữ liệu và không hoạt động với các tệp văn bản riêng lẻ để xây dựng mô hình, như được mô tả trong Deklerck *et al.* (2019).

Kiểm tra mô hình về độ phù hợp và độ chính khi xác phân loại bằng cách loại bỏ phổ tham chiếu (loại bỏ một xác thực chéo, LOOCV) và bằng cách phân tích khối phổ tham chiếu chưa sử dụng

(không sử dụng) trong quá trình xây dựng mô hình (đôi khi được gọi là "khối phổ xác thực"). Các tham số mô hình được tối ưu hóa khi đạt được: (i) LOOCV cao nhất có thể và (ii) phân loại chính xác các khối phổ xác thực.

7.5 Đặt tên file (tệp dữ liệu)

Sử dụng excel để đặt tên cho từng tệp dữ liệu quang phổ khối lượng của các loại gỗ (xem hướng dẫn ở Phụ lục 1).

7.6 Tạo thư viện NIST

Mỗi khối phổ được tìm kiếm trong thư viện ForeST[©] bằng phần mềm Mass Mountaineer. Phần mềm này sử dụng thuật toán do Viện Khoa học và Công nghệ Quốc gia (NIST - National Institute of Science and Technology) phát triển để tìm kiếm trong cơ sở dữ liệu các khối phổ có thành phần hóa học giống nhất với khối phổ của các loại gỗ chưa biết tên.

Sử dụng phần mềm Mass Mountaineer và NIST MS Search để thực hiện các bước tạo thư viện NIST (xem hướng dẫn ở Phụ lục 2).

Thư mục tài liệu tham khảo

[1] Deklerck, V., T. Mortier, N. Goeders, R.B. Cody, W. Waegeman, E. Espinoza, J. Van Acker, J. Van den Bulcke and H. Beeckman (2019). *A protocol for automated timber species identification using metabolome profiling*. Wood Science and Technology.

[2] D., Hawkins (1980). Identification of Outliers. Chapman and Hall, London.

[3] Schmitz, N., Beeckman, H., Blanc-Jolivet, C., Boeschoten, L., Braga, J. W., Cabezas, J. A.,
...& Zuidema, P. A. (2020). Tổng quan về các phương pháp sử dụng trong giám định gỗ. Hướng dẫn về các phương pháp truy xuất gỗ.

[4] US-WISC. Hướng dẫn đặt tên tệp dữ liệu

[5] US-WISC. Hướng dẫn tạo thư viện NIST

[6] US-WISC. Giám định thực vật bằng công nghệ quang phổ khối lượng (DART-TOFMS).

[7] US-WISC. Hướng dẫn phân tích giám định bằng DART TOFMS và Thu thập dữ liệu.

PHỤ LỤC A (Tham khảo) Cách đặt tên tệp dữ liệu

- 1. Để bắt đầu, hãy truy cập cơ sở dữ liệu WD hoặc Ww trong Excel.
- Sao chép các dòng từ Excel của các mẫu đích mà bạn muốn tạo tên tệp bằng cách đánh dấu các dòng rồi nhấn phím Ctrl + C.

Sheet View		Work	cbook Views			Show		Z	oom			
$\overline{\checkmark}$: $\times \checkmark f_x$ Bind	omial_Nor	nenclature	2									
A	В	С	D	E	F	G	Н	I.	J	K	L	М
Binomial_Nomenclature	Ww_Num	Collection	Other_Nu	Subspecie	Sample_	L Previous	Previous	Collector	Source_T	Wild_Cu	lt Specimer	Heartwoo
Abies balsamea	Ww22036	SUNYESF-	8089		Tower 1	Drawer 1	SUNY-ESF		Research		Block	
Abies balsamea	Ww22036	SUNYESF-	8252		Tower 1	Drawer 1	SUNY-ESF		Research		Block	
Abies balsamea	Ww22036	SUNYESF-	8653		Tower 1	Drawer 1	SUNY-ESF		Research		Block	
Abies balsamea	Ww22037	SUNYESF-	8660		Tower 1	Drawer 1	SUNY-ESF		Research		Block	
Abies grandis	Ww21006	Gleaves_:	10_483				Private	William G	Private		Block	
Abies grandis	Ww22037	SUNYESF-	8105		Tower 1	Drawer 1	SUNY-ESF		Research		Block	
Abies grandis	Ww22037	SUNYESF-	8124		Tower 1	Drawer 1	SUNY-ESF		Research		Block	
Abies grandis	Ww22037	SUNYESF-	8366		Tower 1	Drawer 1	SUNY-ESF		Research		Block	
Abies procera	Ww220374	SUNYESF-	8237		Tower 1	Drawer 1	SUNY-ESF		Research		Block	
Abies sp.	Ww21003	TimberEn	gCo_34				Timber Er	gineering	Research		Block	
Acacia auriculiformis	Ww200474	100560					Royal Bot	anical Gard	Research		Block	
Acacia mangium	Ww20001	PZAN256	SOLOMO	N1119	Tower 1	Drawer 1	World For	est ID	Research		Sliver	
Acacia mangium	Ww20001	AOPS164	SOLOMO	v1119	Tower 1	Drawer 1	World For	est ID	Research		Sliver	
Acacia mangium	Ww20002	BHGN301	SIK108		Tower 1	Drawer 1	World For	est ID	Research		Sliver	
Acacia mangium	Ww22037	AAGZ598			Sliver Ca	binet 1 Drav	World For	est ID	Research		Sliver	

3. Chuyển thông tin mẫu đã sao chép sang một trang tính Excel mới bằng cách nhấn phím Ctrl và phím V.

	A	В	С	D	E	F	G	н	1	J	К	L	М	N
1	Binomial_Nomenclature	Ww_Num	Collection_Num	Other_Nums	Subspecie	Sample_Locatic	Previous_	Previous	Collector	Source_T	y Wild_Cult	Specimen	Heartwoo	Notes
2	Abies balsamea	Ww220367	SUNYESF-8089			Tower 1 Drawer	1	SUNY-ESF		Research		Block		
3	Abies balsamea	Ww220368	SUNYESF-8252			Tower 1 Drawer	1	SUNY-ESF		Research		Block		
4	Abies balsamea	Ww220369	SUNYESF-8653			Tower 1 Drawer	1	SUNY-ESF		Research		Block		
5	Abies balsamea	Ww220370	SUNYESF-8660	1		Tower 1 Drawer	1	SUNY-ESF		Research		Block		
6	Abies grandis	Ww210062	Gleaves_10_483					Private	William G	Private		Block		
7	Abies grandis	Ww220371	SUNYESF-8105			Tower 1 Drawer	1	SUNY-ESF		Research		Block		
8	Abies grandis	Ww220372	SUNYESF-8124			Tower 1 Drawer	1	SUNY-ESF		Research		Block		
9	Abies grandis	Ww220373	SUNYESF-8366			Tower 1 Drawer	1	SUNY-ESF		Research		Block		
10	Abies procera	Ww220374	SUNYESF-8237			Tower 1 Drawer	1	SUNY-ESF		Research		Block		
11	Abies sp.	Ww210033	TimberEngCo_34					Timber En	gineering	Research		Block		
12	(Ctrl) -													
13														
1.4														

4. Chỉ giữ lại các cột có nhãn: **Binomial_Nomenclature**, **Ww_Num** hoặc **WD_Num**, **Collection_Num**, **Country**, và **Family**.

А	В	С	D	E	F
Binomial_Nomenclature	Ww_Num	Collection_Num	Country	Family	
Abies balsamea	Ww220367	SUNYESF-8089	New York	Pinaceae	
Abies balsamea	Ww220368	SUNYESF-8252	Minnesota	Pinaceae	
Abies balsamea	Ww220369	SUNYESF-8653	Wisconsin	Pinaceae	
Abies balsamea	Ww220370	SUNYESF-8660	Maine	Pinaceae	
Abies grandis	Ww210062	Gleaves_10_483		Pinaceae	
Abies grandis	Ww220371	SUNYESF-8105	Washington	Pinaceae	
Abies grandis	Ww220372	SUNYESF-8124	Idaho	Pinaceae	
Abies grandis	Ww220373	SUNYESF-8366	Canada	Pinaceae	
Abies procera	Ww220374	SUNYESF-8237	Washington	Pinaceae	
Abies sp.	Ww210033	TimberEngCo_34		Pinaceae	

- 5. Thay đổi cột **Family** thành phiên bản viết tắt của họ bằng tất cả các chữ hoa. Ví dụ:
 - a. Pinaceae PIN
 - b. Fabaceae FAB
 - LƯU Ý: Nhiều Family (gia đình) có 3-4 chữ cái đầu giống nhau, hãy nhớ kiểm tra những chữ viết tắt nào đã được sử dụng trước khi tiếp tục.

Điều này được thực hiện bằng cách đánh dấu cột **Family**, sau đó nhấn phím Ctrl và phím F. Thao tác này sẽ xuất hiện hộp **Find and Replace** (Tìm và Thay thế).

E	F	G	Н	1	J	К	L		M
Family									
Pinaceae	Find a	nd Replace					?	×	
Pinaceae									
Pinaceae	Fing	d Re <u>p</u> lac	e						
Pinaceae	Find	what							
Pinaceae	ri <u>n</u> u	what:							
Pinaceae									
Pinaceae							Op <u>t</u> ions	>>	
Pinaceae								_	
Pinaceae				Find A	II Fin	d Next	Clo	se	
Pinaceae									

Nhấp vào tab **Replace** và nhập họ vào "Find what:" và tên viết tắt trong "Replace with:", sau đó nhấp vào "Replace <u>A</u>ll"

E	F	G	н	1	J	К	L	M					
Family													
Pinaceae	Find an	d Replace					?	X					
Pinaceae													
Pinaceae	Find	Re <u>p</u> la	ice										
Pinaceae	Finds	vhat:	Pinaceae										
Pinaceae		viiat.	Pillaceae	PIN									
Pinaceae	керіа	ce with:	PIN										
Pinaceae						(Op <u>t</u> ions	>>					
Pinaceae													
Pinaceae	Repla	ce <u>A</u> ll	Replace	Find A	All <u>F</u> in	d Next	Clo	se					
Pinaceae													

D	E	F
Country	Family	
New York	PIN	
Minnesota	PIN	
Wisconsin	PIN	
Maine	PIN	
	PIN	
Washington	PIN	
Idaho	PIN	
Canada	PIN	
Washington	PIN	
	PIN	

 Thay đổi cột Binomial_Nomenclature để cả chi và loài đều được viết hoa: sử dụng ô liền kề trong cột F, nhập = PROPER (sau đó nhấp vào cột trống liền kề, đóng dấu ngoặc đơn và nhấn enter.)

	А	В	С	D	E	F	G
I	Binomial_Nomenclature	Ww_Num	Collection_Num	Country	Family		
/	Abies balsamea	Ww220367	SUNYESF-8089	New York	PIN	=PROPER(A2)
1	Abies balsamea	Ww220368	SUNYESF-8252	Minnesota	PIN		
1	Abies balsamea	Ww220369	SUNYESF-8653	Wisconsin	PIN		
1	Abies balsamea	Ww220370	SUNYESF-8660	Maine	PIN		
1	Abies grandis	Ww210062	Gleaves_10_483		PIN		
1	Abies grandis	Ww220371	SUNYESF-8105	Washington	PIN		
1	Abies grandis	Ww220372	SUNYESF-8124	Idaho	PIN		
1	Abies grandis	Ww220373	SUNYESF-8366	Canada	PIN		
1	Abies procera	Ww220374	SUNYESF-8237	Washington	PIN		
1	Abies sp.	Ww210033	TimberEngCo_34		PIN		
	A	В	С	D	E	F	G
1	Binomial_Nomenclature	Ww_Num	Collection_Num	Country	Family		
2	Abies balsamea	Ww220367	SUNYESF-8089	New York	PIN	Abies Bals	amea
3	Abies balsamea	Ww220368	SUNYESF-8252	Minnesota	PIN		
4	Abies balsamea	Ww220369	SUNYESF-8653	Wisconsin	PIN		

 Nhấp và kéo góc dưới cùng bên phải của ô mới được đánh dấu để tự động điền cho tất cả cột Binomial_Nomenclature



8. Nhấp Ctrl + c để sao chép các ô được đánh dấu và sau đó dán chúng dưới dạng giá trị trên các tên **Binomial_Nomenclature** ban đầu. Bằng cách dán dưới dạng giá trị, hàm không bị sao chép, chỉ là tên của loài.

A	$\begin{array}{c cl} Calibn & & \\ \hline \\ \hline$	″o ♥ 😫 .00 🖋	D	E	F G
Binomial_Nomenclat		- ²⁰ 💙 _Num	Country	Family	
Abies balsamea	M/w220367	UNYESF-8089	New York	PIN	Abies Balsamea
Abies balsamea	Search the menus	JNYESF-8252	Minnesota	PIN	Abies Balsamea
Abies balsamea	V c ·	JNYESF-8653	Wisconsin	PIN	Abies Balsamea
Abies balsamea	<u>λ</u> Cu <u>t</u>	JNYESF-8660	Maine	PIN	Abies Balsamea
Abies grandis	[<u>]</u> <u>С</u> ору	eaves_10_483		PIN	Abies Grandis
Abies grandis		JNYESF-8105	Washington	PIN	Abies Grandis
Abies grandis	L Paste Options:	JNYESF-8124	Idaho	PIN	Abies Grandis
Abies grandis		JNYESF-8366	Canada	PIN	Abies Grandis
Abies procera		JNYESF-8237	Washington	PIN	Abies Propera
Abies sp.	Paste <u>Special</u>	mberEngCo_34		PIN	Abies Sp.
	Smart Lookup				

A
Binomial_Nomenclature
Abies Balsamea
Abies Balsamea
Abies Balsamea
Abies Balsamea
Abies Grandis
Abies Grandis
Abies Grandis
Abies Grandis
Abies Procera
Abies Sp.

Cột mới được sử dụng để tạo tên GenusSpecies hiện có thể bị xóa.

9. Thực hiện theo các quy trình tương tự như bước 5, đánh dấu các tên mới, viết hoa trong Binomial_Nomenclature, sau đó nhấn phím cách một lần để tìm khoảng trắng, "" và thay thế bằng không. Điều này sẽ loại bỏ khoảng cách giữa các tên loài.

A	В	С	D	E	F
Binomial_Nomenclature	Ww Num	Collection Num	Country	Family	
Abies Balsamea	Ww Find and Re	place		7	? ×
Abies Balsamea	Ww	Dealers			
Abies Balsamea	Ww Find	Re <u>p</u> lace			
Abies Balsamea	Wv. Find what		Drace tha	enaraha	r 🗸
Abies Grandis	Ww Replace w	ith:			\sim
Abies Grandis	Ww				
Abies Grandis	Ww			Opt	ions >>
Abies Grandis	Ww		nothin	<u> </u>	
Abies Procera	Ww Replace Al	I <u>R</u> eplace F <u>i</u>	nd All <u>F</u> ind	Next	Close
Abies Sp.	Ww210033	TimberEngCo_34		PIN	

A
Binomial_Nomenclature
AbiesBalsamea
AbiesBalsamea
AbiesBalsamea
AbiesBalsamea
Abies Grandis
AbiesGrandis
Abies Grandis
Abies Grandis
AbiesProcera
AbiesSp.

Tất cả các ô hiện đã sẵn sàng để sử dụng làm tên tệp.

10. Trong một cột không sử dụng, hãy nhập = CONCATENATE (và sau đó làm theo định dạng sau:

= CONCATENATE

```
(Family,"_",Binomial_Nomenclature,"_",Ww_Num,"_",Collection_Num,"_",Country)
```

C	$2 \qquad \checkmark \qquad \vdots \qquad \times \checkmark \qquad f_x$	=CONCATENATE(E2,"	",A2,"_",B2,"_",C2,"_",	,D2)						
	A	В	С	D	E	F	G	н	I.	J
1	Binomial_Nomenclature	Ww_Num	Collection_Num	Country	Family					
2	AbiesBalsamea	Ww220367	SUNYESF-8089	New York	PIN	=CONCAT	ENATE(E2,'	"_",A2,"_",	,B2,"_",C2,'	'_",D2)
3	AbiesBalsamea	Ww220368	SUNYESF-8252	Minnesota	PIN					
4	AbiesBalsamea	Ww220369	SUNYESF-8653	Wisconsin	PIN					
5	AbiesBalsamea	Ww220370	SUNYESF-8660	Maine	PIN					
5	AbiesGrandis	Ww210062	Gleaves_10_483		PIN					
7	AbiesGrandis	Ww220371	SUNYESF-8105	Washington	PIN					
В	AbiesGrandis	Ww220372	SUNYESF-8124	Idaho	PIN					
9	AbiesGrandis	Ww220373	SUNYESF-8366	Canada	PIN					
0	AbiesProcera	Ww220374	SUNYESF-8237	Washington	PIN					
1	AbiesSp.	Ww210033	TimberEngCo 34		PIN					

Kết quả sẽ như sau:

PIN_AbiesBalsamea_Ww220367_SUNYESF-8089_New York

Nhấp và kéo góc dưới cùng bên phải của ô mới này giống như trong bước 7 để sao chép chức năng này để tạo tất cả các tên tệp.

PIN_Abies	sBalsamea	_Ww22036	7_SUNYES	F-8089_Ne	w York
PIN_Abies	sBalsamea	_Ww22036	8_SUNYES	F-8252_Mir	nnesota
PIN_Abies	sBalsamea	_Ww22036	9_SUNYES	F-8653_Wi	sconsin
PIN_Abies	sBalsamea	_Ww22037	0_SUNYES	F-8660_Ma	ine
PIN_Abies	sGrandis_V	Vw210062	_Gleaves_1	.0_483_	
PIN_Abies	sGrandis_V	Vw220371	SUNYESF-8	8105_Wasł	nington
PIN_Abies	sGrandis_V	Vw220372	SUNYESF-8	8124_Idaho)
PIN_Abies	sGrandis_V	Vw220373_	SUNYESF-	366_Cana	da
PIN_Abies	sProcera_V	Vw220374_	SUNYESF-	8237_Wash	nington
PIN_Abies	sSpWw2	10033_Tim	berEngCo_	34_	

11. Sử dụng Shift để đánh dấu các tên tệp mới này, nhấn Ctrl + C để sao chép chúng và dán làm Giá trị (giống như bước 8) để thay thế cột đầu tiên trên trang tính.

٨	D	C	D	F	-	C				V		M	N	0
A	Б	C	U	E	F	G	н	1	1	ĸ	L	IVI	IN	0
Binomial_Nomenclature	Ww_Num	Collection_Num	Country	Family										
PIN_AbiesBalsamea_Ww220367_SU	Ww220367	SUNYESF-8089	New York	PIN	PIN_PIN_	AbiesBalsa	mea_Ww2	220367_SU	NYESF-8089	_New Yor	_Ww2203	67_SUNYE	SF-8089_N	ew York
PIN_AbiesBalsamea_Ww220368_SU	Ww220368	SUNYESF-8252	Minnesota	PIN	PIN_PIN_	AbiesBalsa	mea_Ww2	220368_SU	NYESF-8252	_Minneso	ta_Ww220	368_SUNY	ESF-8252_I	Minnesota
PIN_AbiesBalsamea_Ww220369_SU	Ww220369	SUNYESF-8653	Wisconsin	PIN	PIN_PIN_	AbiesBalsa	mea_Ww2	220369_SU	NYESF-8653	_Wisconsi	n_Ww220	69_SUNYE	SF-8653_V	Visconsin
PIN_AbiesBalsamea_Ww220370_SU	Ww220370	SUNYESF-8660	Maine	PIN	PIN_PIN_	AbiesBalsa	mea_Ww2	220370_SU	NYESF-8660	_Maine_W	/w220370_	SUNYESF-	3660_Main	e
PIN_AbiesGrandis_Ww210062_Glea	Ww210062	Gleaves_10_483		PIN	PIN_PIN_	AbiesGran	dis_Ww21	0062_Glea	ves_10_483	Ww2100	62_Gleave	es_10_483	-	
PIN_AbiesGrandis_Ww220371_SUN	Ww220371	SUNYESF-8105	Washington	PIN	PIN_PIN_	AbiesGran	dis_Ww22	0371_SUN	/ESF-8105_\	Nashingto	n_Ww2203	71_SUNYE	SF-8105_V	Vashington
PIN_AbiesGrandis_Ww220372_SUN	Ww220372	SUNYESF-8124	Idaho	PIN	PIN_PIN_	AbiesGran	dis_Ww22	0372_SUN	/ESF-8124_I	daho_Ww	220372_SU	NYESF-812	4_Idaho	
PIN_AbiesGrandis_Ww220373_SUN	Ww220373	SUNYESF-8366	Canada	PIN	PIN_PIN_	AbiesGran	dis_Ww22	0373_SUN	/ESF-8366_0	Canada_W	w220373_9	UNYESF-8	366_Canad	а
PIN_AbiesProcera_Ww220374_SUN	Ww220374	SUNYESF-8237	Washington	PIN	PIN_PIN_	AbiesProce	era_Ww22	0374_SUN	(ESF-8237_)	Nashingto	n_Ww2203	74_SUNYE	SF-8237_V	Vashington
PIN_AbiesSpWw210033_TimberEr	Ww210033	TimberEngCo_34		PIN	PIN PIN	AbiesSp\	Ww210033	TimberEn	gCo_34V	/w210033_	TimberEn	gCo_34_		

Tất cả các dữ liệu khác có thể bị xóa ngoại trừ các tên tệp mới được tạo.

12. Nếu có bất kỳ tên tệp nào không có quốc gia được liệt kê trong cơ sở dữ liệu, sẽ có thêm dấu "_" ở cuối tên tệp được tạo. Chúng có thể bị xóa để để lại phần cuối cùng của tên tệp là Collection_Num.

File Names
PIN_AbiesBalsamea_Ww220367_SUNYESF-8089_New York
PIN_AbiesBalsamea_Ww220368_SUNYESF-8252_Minnesota
PIN_AbiesBalsamea_Ww220369_SUNYESF-8653_Wisconsin
PIN_AbiesBalsamea_Ww220370_SUNYESF-8660_Maine
PIN_AbiesGrandis_Ww210062_Gleaves_10_48
PIN_AbiesGrandis_Ww220371_SUNYESF-8105_Washington
PIN_AbiesGrandis_Ww220372_SUNYESF-8124_Idaho
PIN_AbiesGrandis_Ww220373_SUNYESF-8366_Canada
PIN_AbiesProcera_Ww220374_SUNYESF-8237_Washington
PIN_AbiesSpWw210033_TimberEngCo_34

Tên cuối cùng sẽ như sau:

1	File Names
2	PIN_AbiesBalsamea_Ww220367_SUNYESF-8089_New York
3	PIN_AbiesBalsamea_Ww220368_SUNYESF-8252_Minnesota
4	PIN_AbiesBalsamea_Ww220369_SUNYESF-8653_Wisconsin
5	PIN_AbiesBalsamea_Ww220370_SUNYESF-8660_Maine
6	PIN_AbiesGrandis_Ww210062_Gleaves_10_483
7	PIN_AbiesGrandis_Ww220371_SUNYESF-8105_Washington
8	PIN_AbiesGrandis_Ww220372_SUNYESF-8124_Idaho
9	PIN_AbiesGrandis_Ww220373_SUNYESF-8366_Canada
10	PIN_AbiesProcera_Ww220374_SUNYESF-8237_Washington
11	PIN_AbiesSpWw210033_TimberEngCo_34

PHŲ LỤC B

(Tham khảo)

Cách tạo thư viện NIST

1. Mở phần mềm Mass Mountaineer và chọn thanh công cụ Spectrum



a. LƯU Ý: Ngưỡng mặc định cho quy trình này là 0,5, nếu cần phải điều chỉnh ngưỡng, bạn có thể thay đổi bằng cách điều hướng đến nút Edit trong Spectrum.
 Mass Mountaineer

👷 Spectrum 🚺	l, Con	nposition	lsoto	pes 📶	Series	Beo	MS Periodi	c Table	Classify	ESI	P
File Profile	Edit	View	Options	Print to	Word or	RTB	Tools	Search	Search	multip	le chi
View Spectra		Сору			•	d m/:	z difference	s Comp	pare Spec	tra N	latch
		Threshold	for batch	processii	ng 🕨		0.5				
Mass Spe		Attenuate	selected p	eak by	+	SW	ap spectra				
Compariso	on MS					<u> </u>					

2. Đặt đường dẫn đến thư mục khối phổ sẽ được sử dụng để tạo thư viện tìm kiếm NIST.

🕍 Ma	ass Mountaineer						
Ωs	Spectrum 👯 Composition	lsotopes	🛗 Series	6. N	/IS Period	dic Table	Clas
File	e Profile Edit <mark>View</mark>	Options Prir	nt to Word or	RTB	Tools	Search	Sea
	Open		•	nd m/z	differenc	es Comp	bare :
	Set default directory for ma	ss spectra					
	Set default directory for con	npound lists		Swa	p spectra	э	
	Get MS from Clipboard		Ctrl+G			_	_
	Get last MS that was sent to	NIST	Ctrl+N		`	Clear	
	Clear Search List						
	Printer Properties		r	rance ((mmu)		
	Print Now		Ctrl+P	5			
	Print Now Goes to Printer (D	oes not use W	/ord)		- 0/ □	Deletivet	_
	Print to PDF (through Word)		-	eshold		Relative to	- -
	Printer Preview		ŀ	5	B	ase peak	
	Print to Word File	C	Ctrl+W D)imers,	, Trimers		
	Save MS in NIST MSP Form	at					
	Save MS in Mascot MS/MS f	format	e	earch	Fo	ound	
	Save Search List as Excel Fi	le	ł		Abur	idance	
	Save search List as CSV (te	xt) file		lear		0	
	Save search list masses +/-	tolerances	l	dd to o	visting se	arch	
	Save mass differerence ann	otations to file		ark leo	tone Mat	tchos (*)	
	Read mass difference annot	ations from fil	e	nlv sho	w isotop	e matches	
	Save defaults file now			,		e materies	
	Exit		Ctrl+X	Searc	h Modific	ations	
	Clear Comparison MS	Results to DE	3 Search	Мо	dification	List	
		Delete Searc	h Peaks				

3. Chọn Batch Processing và chọn Export to NIST MSP. Sau đó bỏ chọn bất kỳ tùy chọn nào khác đã chọn .



4. Chọn Batch process để xử lý hàng loạt tất cả các tệp trong thư mục

es Ba	atch processing Help	_	-		
ectr	Just Print Centroided Mass Spectra				
	Compound search	_		Calophyllaceae	^
	Compositions only			Caryocaraceae	
	Composition and isotope match			Clusiascase	
	Search for matching spectra			Combretaceae	
	Print Mass Defect Graph to Word			Cunoniaceae	
~	Export to NIST MSP			Dilleniaceae	
	Export KMD analysis to Excel			Dipterocarpaceae	
				Ebenaceae	
	I hreshold before processing			Ericaceae	
	Centroid before processing			Erythroxylaceae	
	Save centroided text files			Euphorbiaceae	
	Purge Compound List masses from spectrum		=	Fagaceae	
	Save 3 column text files			Gentianaceae	
	Truncate mass range only			Irvingiaceae	
_	GC: BT to BI	_		Juglandaceae	
		_		Lamiaceae	
	Batch process all files in directory			Lauraceae	
	Batch process selected files			Lecythidaceae	
	Limit m/z range for batch report			Magnoliaceae	
	Lower m/z limit			Malyaceae	
	Upper m/z limit			> Meliaceae	
	Interval +				~
	Batch Search and Match		Make New Folder		OK Cancel
	Summary_only				
	Combine Integer Mass Spectra				

5. Thay đổi Decimal places for m/z values thành 4. Bỏ chọn hộp MS/MS. Tất cả các thiết lập khác phải khớp với hình ảnh bên dưới. Nhấp vào OK và kiểm tra xem các tệp đã xuất của bạn sẽ được lưu trong đúng đường dẫn.

H.	🖳 Spectrum details	_	П	×	11		
	- opectrum details			~		Mass_Mountaineer_v6	Х
	Title	FAB_XyliaXylocarpa_WISC	w20064	6_Kv			
F	Comment	JEOL USA, Inc. Den	no Lab			The exported NIST-format files will be saved	
	Instrument	JEOL SpiralTO	F			inZ:\Erin\ForeST_2020\Angiosperms\FabaceaeNISTout.MSP	
_	Instrument type	TOF/TOF					
3	Decim	nal places for m/z value	s 4	~		ОК]
		MS/MS					
		Omit precursor peak					
٩	Pre	cursor m/z 2					
	Prec	ursor type [M + H]+					
l	Collisi	on energy 20 kV					
v		Ionization MALDI					
	Co	Ilision gas He	~				
			OK				

6. Có thể theo dõi quá trình xử lý ở góc dưới bên trái của màn hình Mass Mountaineer. Các tập dữ liệu lớn sẽ mất nhiều thời gian hơn.

<u> </u>	
Progress	Status:

7. Kiểm tra thư mục khối phổ để tìm tệp sau:

Name	✓ Date modified	Туре	Size
侵 NISTout	2/1/2022 2:34 PM	Windows Installer Patch	62,780 KB
8. Mở MS Search	và chọn Librariar	n ở phía dưới bên trái của	màn hình

<i>p</i>						
Lib. Search	Other Search	Names	Compare	Librarian	MSMS	Γ

9. Chọn Options ở phía trên và chọn Spectrum Import Options

$ \begin{array}{c} & & & & & & & & & & & & & & & & & & &$	File Search View Tools	Options Window Help	
Spectrum Import Options TiSpec. List) Caffeine Pubchem search options SOS options OH HIN HIN HIN HO HO<	, h t .	^{ar/z} m/z range Library <u>S</u> earch Options Replicates	
I (Spec. List) Catterie Pubchem search options Ust) Xanthine SOS options OH HIN HIN HIN HO Name Specific Specif		Spectrum Import Options	
SOS options HN HN HN HN HN HO Number (1 (Spec. List) Caffeine	Pubchem search options	List) Xanthine
Names Structures Special		SOS options	
January Astronomy Speech	Names A Structures		Spec List

10. Chọn Accurate m/z. Tất cả các thiết lập khác phải như sau:
 a. LƯU Ý: Không sử dụng cài đặt nominal m/z.

Spectrum Import Options	×
EI Spectrum m/z Rounding Multiply m/z in imported spectra by and round to the nearest integer	Tandem Spectrum Accuracy Precursor ion m/z 4 v decimal places
Example: (CH2)n correction is 0.99888 Optional Add this term to all m/z	Product ion m/z 4 v decimal places
0 before rounding	In-source/EI accurate ion m/z
Adding spectra to Spec List Prepend (add to the top) Overwrite (replace) Ask	Intensity threshold % of max. 0 absolute 0
Spectra without precursor ion m/z va Accurate EI (nominal m/z)	alue are e m/z spectrum type urce
RI type if unspecified	Unspecified V
Include Synonyms OK Ca	Set default values

11. Chọn biểu tượng Import ở phía trên bên trái của màn hình.



12. Tìm tệp dữ liệu NISTout được tạo từ Bước 6. Sau đó chọn Import All. Hủy tìm kiếm Background, thao tác này tự động nhưng không cần thiết và sẽ làm tăng thời gian cho quá trình xử lý.

				Importing spectra	×
N	umbei	r of Spectra found: 4440	×	FAB_MicroberliniaBisulcata_WD172606-Light_Tw22582-2.txt	Cancel
	#	The names of spectra retrieved	Import All	29%	
	1	FAB_AcaciaAuriculiformis_WISCw200474_Kw1	Impart Calastad		
	2	FAB_AcaciaBakeri_WD140658_USGSR-1997.	import Selected		
	3	FAB_AcaciaHarmandiana_WISCw200475_Kw.	Import Options		
	4	FAB_AcaciaHeterophylla_WD173651_Drouin.t:			
	5	FAB_AcaciaHeterophylla_WD173652_Drouin.t:	Search Options		
	6	FAB_AcaciaHeterophylla_WD173653_Drouin.t:	Consul		
	7	FAB_AcaciaHeterophylla_WD173654_Drouin.t:	Cancel		
	2		Help		
	<	>	neip		

13. Bây giờ các khối phổ đã được nhập, chọn Create Library



14. Nhập tên thư viện ngắn gọn và đơn giản. Sau đó nhấp vào OK Create library ×

Enter new library nam	e	_	ОК
List of libraries			Cancel
mainlib replib angio-2021 april 8 angiobrazil-2021 focusedforest-21 forest_2021 forest_gymno2021	~		Help
Library Statistics			
242466	Spectra		
1 - 242469	ID		

15. Nhấp vào bất kỳ tệp khối phổ nào trong ô bên trái (Tên) rồi giữ Ctrl + A để chọn tất cả. Nhấp vào biểu tượng Di chuyển đến Thư viện. Nhấp vào OK.



16. Chọn tên thư viện vừa tạo và nhấp vào OK. Thư viện sẽ bắt đầu tích hợp các quang phổ.

	205.1594 Moving spectra	/	1
11	Moving to the library: "africa_cm" MOR_MiliciaRegia_WD166842_Tw21224_c 11%	Cancel leaned.txt	0
B ra	te m/z	[©]	

17. Sau bước cuối cùng này, thư viện đã sẵn sàng để sử dụng. Có thể thay đổi thư viện bằng cách điều hướng đến tab Composition, chọn NIST Search và chọn thư viện mục tiêu từ danh sách.

NIST Entrice for Matching compounds		Copy Sele	ected Name
NIST Entres for Matching compounds	[Copy Selec	cted Formu
Name	Formula	F Match	R Match
FAB_DalbergiaChlorocarpa_WD21170		833	833
FAB_DalbergiaGreveana_WD211694		823	824
FAB_DalbergiaGreveana_WD210984		823	823
FAB_DalbergiaNeoperrieri-AFF_WD2		821	821
FAB_DalbergiaHildebrandtii_WD2109		820	820
FAB_DalbergiaNeoperrieri-AFF_WD2		818	818
FAB_DalbergiaNeoperrieri-AFF_WD2		817	817
FAB_DalbergiaGreveana_WD211234		815	815
FAB_DalbergiaBemarivensis_WD2109		812	812
FAB_DalbergiaChlorocarpa_WD21130		811	811
FAB_DalbergiaNeoperrieri-AFF_WD2		811	812
FAB_DalbergiaPurpurascens-CF_WD		811	813
FAB_DalbergiaChlorocarpa_WD21123		810	810
Search Formula 01	Search Na	ime	
СЗН6О	Benzene	e	
Search	Search	🗹 a-z on	y
Print To Word File	Isotope Gra	ph	
	mainlib	^	
Detabased	Africa_CM	dueed	
Press CTRL to select > 1	AngioBRAZIL	-2021	
	Angio 2022	~	